수리통계학 × 통계학의 역사

11회 워크북

충분성
손으로 푸는 34문항

강의노트 11회를 읽은 뒤 손으로 푼다. 인수분해, Rao–Blackwell, Lehmann–Scheffé, Basu의 증명은 한 줄씩 짧지만, 각 줄에 쓰인 근거를 빈 종이에서 다시 찾는 것이 이 워크북의 일이다. 문제 아래의 줄 친 공간에 직접 쓴다.

부	하는 일	문항	예상 시간
A 계산 드릴	조건부분포, UMVUE, 로그분배함수, 정보를 손으로 계산한다	14	2시간 4분
B 빈칸 증명	노트의 단계형 증명을 책을 덮고 다시 쓴다	7	2시간 40분
C 증명 문제	변형, 일반화, 반례 만들기	9	2시간 50분
D 퀄형	여러 절을 잇는 종합 문제. 제한 시간 안에	4	2시간 20분
	합계	34	약 10시간

노트 `11_sufficiency_notes.pdf` — 문항의 "정리 2.1" 같은 번호는 노트의 번호다

도구 `00_toolkit_notes.pdf` — [T1] 원소 추적, [T2] 귀류·반례, [T4] 부등식, [T6] 극한의 교환, [T7] 변수변환, [T8] 델타 방법

검산 노트북 `11_sufficiency.ipynb`의 "워크북 검산" 절: `check("A3a", "2/15")`

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. 식에 λ가 들어가면 `lam`으로 쓴다(`lambda`는 파이썬의 예약어다).
5. **증명은 근거를 단다.** 등호와 부등호마다 어느 정의, 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.

A

계산 드릴

정의와 공식을 손으로 돌린다. 답이 수나 식으로 떨어지므로 다 풀면 뒤의 최종답이나 노트북의 `check()`로 맞춰 본다. 계산기는 마지막 수치 계산에만 쓴다.

A1 조건부분포에서 모수가 사라진다 ★

정의 1.1, 예제 1.2 · 8분

(a) $X_1, \dots, X_5 \overset{\text{iid}}{\sim} \text{Bernoulli}(p)$, $T = \sum X_i$이다. $P(\mathbf{X} = (1,1,0,1,0) \mid T = 3)$을 구하고 p가 사라짐을 확인하라. (b) $X_1, X_2, X_3 \overset{\text{iid}}{\sim} \text{Poi}(\lambda)$, $T = \sum X_i$이다. $P(\mathbf{X} = (2,1,1) \mid T = 4)$를 구하라.

A2 Poisson 표본의 UMVUE 값 ★

정리 2.1, 3.3 · 8분

$X_1, \dots, X_4 \overset{\text{iid}}{\sim} \text{Poi}(\lambda)$에서 $T = \sum X_i = 6$을 관측했다. (a) $e^{-\lambda} = P(X = 0)$의 UMVUE의 값을 구하라. (b) $P(X = 2) = \frac{1}{2}\lambda^2 e^{-\lambda}$의 UMVUE를 $1\{X_1 = 2\}$에 T로 조건을 걸어 구하고, 그 값을 계산하라.

A3 Bernoulli의 UMVUE ★

예제 2.4, 정리 3.3 · 6분

$X_1, \dots, X_{10} \overset{\text{iid}}{\sim} \text{Bernoulli}(p)$에서 $T = \sum X_i = 4$를 관측했다. (a) p^2의 UMVUE의 값, (b) $p(1-p)$의 UMVUE의 값을 구하라. 각각을 MLE $\hat{p}^2 = 0.16$, $\hat{p}(1-\hat{p}) = 0.24$와 비교하라.

A4 균등분포의 UMVUE ★★

예제 3.2(c), 정리 3.3 · 8분

$X_1, \dots, X_4 \overset{\text{iid}}{\sim} \text{U}(0, \theta)$에서 1.2, 3.1, 0.7, 2.5를 얻었다. (a) θ의 UMVUE의 값을 구하라. (b) $1/\theta$의 UMVUE를 $c/X_{(n)}$ 꼴에서 찾고, 그 값을 구하라.

A5 독일 전차 ★

예제 2.3, 3.4(d) · 8분

$\{1, \dots, N\}$에서 비복원으로 뽑은 일련번호 19, 40, 42, 60을 얻었다. (a) N의 UMVUE의 값과 (b) 적률법 추정값 $2\bar{X} - 1$을 구하라. (a)가 (b)를 Rao–Blackwell화한 것임을 한 줄로 말하라.

A6 기하분포의 로그분배함수 ★★

정의 4.1, 정리 4.3 · 12분

$P(X=x)=p(1-p)^{x-1}$ $(x=1,2,\dots)$를 자연모수 $\eta=\log(1-p)$, $T(x)=x$, $h(x)=1$인 지수족으로 쓴다. (a) $A(\eta)$를 구하라. (b) $A'(\eta)$로 $p=\frac{1}{4}$에서 EX를, (c) $A''(\eta)$로 $\operatorname{Var}X$를 구하라.

A7 두 모수화의 정보 ★

따름정리 4.4 · 8분

Bernoulli(p), $p=0.2$에서 (a) 자연모수 $\eta=\log\frac{p}{1-p}$에 대한 한 관측의 정보 $I(\eta)$와 (b) p에 대한 정보 $I(p)$를 구하라. 두 값의 관계를 $\frac{d\eta}{dp}$로 설명하라.

A8 지수 표본의 최댓값의 몫 ★★

정리 5.2 · 10분

$X_1,\dots,X_4 \overset{\text{iid}}{\sim}$ Exp(평균 θ)이다. $EX_{(n)}=\theta\left(1+\frac{1}{2}+\cdots+\frac{1}{n}\right)$을 써서 $E\left[X_{(4)}/\sum_i X_i\right]$를 구하라.

A9 균등 표본의 순서통계량의 몫 ★

예제 5.5(b) · 8분

$X_1, \dots, X_5 \overset{\text{iid}}{\sim} \mathrm{U}(0,\theta)$이다. $EX_{(k)} = \frac{k}{n+1}\theta$를 써서 $E\left[X_{(2)}/X_{(5)}\right]$를 구하라.

A10 UMVUE의 분산과 Cramér–Rao 한계 ★★

정리 2.1, 9회 정리 2.4 · 10분

$X_1, \dots, X_{10} \overset{\text{iid}}{\sim} \mathrm{Poi}(\lambda)$, $\lambda = 1$에서 $e^{-\lambda}$의 UMVUE $(1-\frac{1}{n})^T$를 생각한다. 확률생성함수 $Es^T = e^{n\lambda(s-1)}$을 써서 (a) $\mathrm{Var}\big((1-\frac{1}{n})^T\big)$, (b) Cramér–Rao 한계 $\frac{\tau'(\lambda)^2}{nI(\lambda)}$, (c) 둘의 비를 구하라.

A11 모수가 두 개뿐이면 완비가 아니다 ★★

정의 3.1 · [T2] · 10분

$X_1, X_2 \overset{\text{iid}}{\sim} \mathrm{Bernoulli}(p)$이고 모수공간이 $p \in \{\frac{1}{4}, \frac{3}{4}\}$ 두 점뿐이다. $T = X_1 + X_2$의 함수 g 가운데 두 p에서 모두 $E_p g(T) = 0$인 것을 $g(0) = 1$로 정규화해 찾아라. (a) $g(1)$과 (b) $g(2)$를 구하라.

A12 Rayleigh 분포의 MLE ★★

정리 4.3, 따름정리 4.4 · 8분

Rayleigh 밀도 $f(x;\theta)=\frac{x}{\theta}e^{-x^2/(2\theta)}$ $(x>0)$를 $T(x)=x^2$인 지수족으로 쓰고, $A'(\eta)=EX^2$를 구하라. 자료 $1,\ 2,\ 2,\ 3$에서 가능도 방정식 "충분통계량의 평균 = 모형의 평균"으로 $\hat\theta$를 구하라.

A13 오염 정규에서 평균편차의 효율 ★

명제 7.1 · 10분

$X\sim(1-\varepsilon)\mathcal{N}(0,1)+\varepsilon\mathcal{N}(0,9)$, $\varepsilon=0.02$에서 μ_2, μ_4, ν_1을 구하고, 명제 7.1의 식으로 평균편차의 상대효율을 소수 셋째 자리까지 구하라.

A14 "0인가 아닌가"가 담는 정보 ★★

명제 1.7 · 10분

$X\sim\mathrm{Poi}(\lambda)$ 하나에서 $T=\mathbb{1}\{X=0\}$만 기록한다. (a) T의 정보 $I_T(\lambda)$를 구하라(답에는 λ를 `lam`으로 쓴다). (b) $\lambda=2$에서 I_T/I_X를 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 정의나 정리의 번호, 앞 회차의 정리, 도구 번호를 쓴다.

B1 인수분해 정리, 근거 채우기 ★★

① · 정리 1.3 · 20분

이산 표본에서 T가 충분통계량일 필요충분조건은 $f(\mathbf{x};\theta) = g(T(\mathbf{x});\theta)h(\mathbf{x})$인 것이다. 각 단계의 근거를 써라.

(⇒)

① $t = T(\mathbf{x})$이면 $\{\mathbf{X} = \mathbf{x}\} \subset \{T = t\}$이다. 왜? _____

② $P_\theta(T = t) > 0$이면 $f(\mathbf{x};\theta) = P_\theta(T = t)\,P_\theta(\mathbf{X} = \mathbf{x} \mid T = t)$이다.

왜? _____

③ 둘째 인수 $h(\mathbf{x}) = P(\mathbf{X} = \mathbf{x} \mid T = T(\mathbf{x}))$는 θ에 의존하지 않는다.

왜? _____

④ $g(t;\theta) = P_\theta(T = t)$로 두면 $f(\mathbf{x};\theta) = g(T(\mathbf{x});\theta)\,h(\mathbf{x})$이다. 왜? _____

(⇐)

① $T(\mathbf{x}) = t$, $P_\theta(T = t) > 0$이면 $P_\theta(T = t) = g(t;\theta)\sum_{\mathbf{y}:\,T(\mathbf{y})=t} h(\mathbf{y})$이다.

왜? _____

② $P_\theta(\mathbf{X} = \mathbf{x} \mid T = t) = \dfrac{g(t;\theta)h(\mathbf{x})}{g(t;\theta)\sum_{\mathbf{y}} h(\mathbf{y})} = \dfrac{h(\mathbf{x})}{\sum_{\mathbf{y}:\,T(\mathbf{y})=t} h(\mathbf{y})}$이다.

왜? _____

③ 오른쪽에 θ가 없으므로 T는 충분하다. 왜? _____

B2 Rao–Blackwell, 근거 채우기 ★★

① · 정리 2.1 · 15분

W가 $\tau(\theta)$의 불편추정량이고 $\mathrm{Var}_\theta\, W < \infty$, T가 충분통계량이면 $\varphi(T) = E(W \mid T)$는 통계량이고 불편이며 $\mathrm{Var}\,\varphi(T) \le \mathrm{Var}\, W$이다. 근거를 채워라.

① $T = t$가 주어질 때 $\mathbf{X}$의 조건부분포는 θ에 의존하지 않는다.

왜? _____

② $\varphi(t) = E(w(\mathbf{X}) \mid T = t)$는 θ 없이 계산된다. 왜? _____

③ $E_\theta \varphi(T) = E_\theta W = \tau(\theta)$이다. 왜? _____

④ $\mathrm{Var}_\theta\, W = \mathrm{Var}_\theta\, \varphi(T) + E_\theta\, \mathrm{Var}(W \mid T)$이다. 왜? _____

⑤ $\mathrm{Var}_\theta\, \varphi(T) \le \mathrm{Var}_\theta\, W$이다. 왜? _____

⑥ 등호이면 $E_\theta(W - \varphi(T))^2 = 0$, 곧 $W = \varphi(T)$ (확률 1)이다. 왜? _____

B3 Basu의 정리, 근거 채우기 ★★

① · 정리 5.2 · 15분

T가 완비충분통계량이고 V가 보조통계량이면 $T \perp V$이다. 근거를 채워라.

① $p_B = P_\theta(V \in B)$는 θ에 의존하지 않는다. 왜? _____

② $\eta(T) = P(V \in B \mid T)$는 θ를 포함하지 않는 T의 함수다. 왜? _____

③ 모든 θ에서 $E_\theta[\eta(T) - p_B] = 0$이다. 왜? _____

④ 모든 θ에서 $\eta(T) = p_B$ (확률 1)이다. 왜? _____

⑤ $P_\theta(T \in C, V \in B) = E_\theta[1\{T \in C\}\eta(T)] = p_B P_\theta(T \in C)$이다.

왜? _____

B4 인수분해 정리, 단계의 이름만 ★★

② · 정리 1.3 · 25분

이산 표본에서 정리 1.3의 두 방향을 증명하라. 단계의 이름만 주어진다. 각 단계 아래에 식과 근거를 써라.

① **(⇒) $\{\mathbf{X} = \mathbf{x}\}$와 $\{T = t\}$의 관계**

② **결합확률을 곱셈법칙으로 쪼갠다**

③ **충분성으로 h를 정하고 g를 정한다**

④ **(⇐) $P_\theta(T=t)$를 g와 h로 쓴다**

⑤ **조건부확률을 계산하고 g를 약분한다**

⑥ **$P_\theta(T=t)=0$인 경우를 처리한다**

B5 Rao–Blackwell, 단계의 이름만 ★★

② · 정리 2.1 · 25분

정리 2.1의 (a)–(c)를 증명하라.

① **(a) 충분성으로 φ가 통계량임을 보인다**

② **(b) 반복기댓값으로 불편성을 보인다**

③ **(c) 전분산 공식을 쓴다**

④ **조건부분산이 음이 아님을 쓴다**

⑤ **등호 조건을 분석한다**

B6 인수분해 정리, 전략만 ★★

③ · 정리 1.3 · 30분

이산 표본에서 정리 1.3을 증명하라.

전략 (⇒)는 결합확률을 "T의 확률 × T가 주어진 조건부확률"로 쪼갠다. (⇐)는 조건부확률을 계산해 g를 약분한다.

B7 Basu의 정리, 전략만 ★★

③ · 정리 5.2 · 30분

정리 5.2를 증명하라.

전략 사건 $\{V \in B\}$의 조건부확률 $\eta(T)$를 만든다. 그 평균이 모든 θ에서 같은 상수이므로 완비성으로 $\eta(T)$ 자체가 그 상수다.

C

증명 문제

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

C1 충분하지 않은 통계량에 조건을 걸면 ★ [T2] · 정리 2.1 · 15분

$X_1, X_2 \overset{\text{iid}}{\sim} \mathcal{N}(\theta, 1)$이고 $W = \bar{X}$이다.

(a) $T = X_1$이 θ의 충분통계량이 아님을 X_2의 조건부분포로 보여라.

(b) $E(W \mid X_1)$을 구하고, 그것이 추정량이 아닌 이유를 말하라.

(c) $\operatorname{Var} E(W \mid X_1) < \operatorname{Var} W$임을 확인하라. 정리 2.1의 어느 부분이 성립하고 어느 부분이 깨지는가?

C2 굽은 정규족 ★★

[T2] · 정리 1.6, 4.5 · 20분

$X_1, \dots, X_n \overset{\text{iid}}{\sim} \mathcal{N}(\theta, \theta^2)$, $\theta > 0$, $n \geq 2$이다.

(a) $\left(\sum X_i,\ \sum X_i^2\right)$가 최소충분통계량임을 정리 1.6으로 보여라. 모수는 하나인데 통계량은 2차원이다.

(b) $E\bar{X}^2$과 ES^2을 구하고, 모든 θ에서 평균이 0이지만 0이 아닌 함수 $g(\bar{X}, S^2)$을 찾아 이 통계량이 완비가 아님을 보여라.

(c) 정리 4.5가 여기에 쓰이지 않는 이유를 자연모수의 모양으로 설명하라.

C3 이동한 지수분포의 UMVUE ★★

[T6] · 정리 1.3, 3.3 · 20분

$X_1, \dots, X_n \overset{\text{iid}}{\sim} f(x;\theta) = e^{-(x-\theta)}\ (x > \theta)$이다.

(a) $X_{(1)}$이 충분통계량임을 인수분해 정리로 보여라.

(b) $P(X_{(1)} - \theta > u) = e^{-nu}\ (u > 0)$임을 보여라.

(c) $X_{(1)}$이 완비임을 보여라.

(d) θ의 UMVUE를 구하고, $n = 5$, $X_{(1)} = 2.3$일 때 그 값을 구하라.

C4 치우친 추정량의 Rao–Blackwell화 ★ 정리 2.1 · 15분

W가 $\tau(\theta)$의 추정량이고(불편일 필요는 없다) $\mathrm{Var}_\theta\, W < \infty$, T가 충분통계량이다. $\varphi(T) = E(W \mid T)$이면 모든 θ에서

$$E_\theta\big(\varphi(T) - \tau(\theta)\big)^2 \le E_\theta\big(W - \tau(\theta)\big)^2$$

임을 보여라. 평균제곱오차를 분산과 치우침으로 나누어 설명하라.

C5 정규 표본의 범위 ★★

정리 5.2 · 20분

$X_1, \dots, X_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$이고 $R = X_{(n)} - X_{(1)}$은 범위다.

(a) σ를 고정하고 $\bar{X} \perp R$임을 보여라.

(b) 두 모수가 모두 미지일 때 R/S가 $(\bar{X}, S)$와 독립임을 보이고, $E(R/S) = ER/ES$를 얻어라.

(c) R과 S는 독립이 아니다. $\operatorname{Cov}(R, S) = E(R/S)\operatorname{Var} S > 0$을 보여라.

C6 충분통계량의 함수 ★★

[T2] · 정리 1.3 · 15분

(a) T가 충분하고 어떤 통계량 U와 함수 r로 $T = r(U)$이면 U도 충분함을 보여라.

(b) 충분통계량의 일대일 함수는 충분함을 보여라.

(c) 반례: $X_i \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$에서 두 모수가 모두 미지이면 $\bar{X}$ 하나는 충분하지 않음을 $\operatorname{Var}(X_1 \mid \bar{X})$로 보여라.

C7 평균모수와 Cramér–Rao 한계 ★★

정리 4.3, 4.6 · 20분

1-모수 지수족 $f(x;\eta)=h(x)\exp\big(\eta T(x)-A(\eta)\big)$에서 $\operatorname{Var}_\eta T>0$이고, 평균모수를 $\mu=A'(\eta)$라 하자.

(a) μ가 η의 순증가함수임을 보여라.

(b) μ에 대한 한 관측의 정보가 $I(\mu)=1/\operatorname{Var} T$임을 보여라.

(c) $\bar T=\frac{1}{n}\sum T(X_i)$가 μ의 불편추정량이고 분산이 Cramér–Rao 한계와 같음을 보여라. 정리 4.6과 어떻게 맞는가?

C8 UMVUE는 0의 불편추정량과 상관이 없다 ★★★

[T4] · 정리 3.3 · 20분

W는 $\tau(\theta)$의 불편추정량이고 분산이 유한하다. $\mathcal{U}$를 0의 불편추정량($E_\theta U=0$, 모든 θ) 가운데 분산이 유한한 것의 모임이라 하자.

(a) W가 UMVUE이면 모든 $U\in\mathcal{U}$와 θ에서 $\operatorname{Cov}_\theta(W,U)=0$임을 보여라.

(b) 거꾸로 모든 $U\in\mathcal{U}$와 θ에서 $\operatorname{Cov}_\theta(W,U)=0$이면 W가 UMVUE임을 보여라.

(c) 완비충분통계량 T가 있으면 $\mathcal{U}$에서 T의 함수는 0뿐이다. 이것과 (b)가 Lehmann–Scheffé 정리와 어떻게 이어지는지 말하라.

C9 정보를 보존하면 충분하다 ★★★ [T6] · 명제 1.7 · 25분

이산 표본이고 9회의 정칙조건이 성립하며, 모수공간 Ω는 구간이다. 모든 θ에서 $I_T(\theta) = I_{\mathbf{X}}(\theta) < \infty$이면 T가 충분함을 보여라. (명제 1.7의 역)

D

퀄형

여러 절의 도구를 잇는 종합 문제다. 머리의 제한 시간을 재고, 노트를 덮고 푼다. UMVUE를 구하는 문제는 "완비충분통계량 찾기 → 불편추정량 하나 → 조건부기댓값" 의 순서를 답안에 드러낸다.

D1 지수분포의 생존확률 ★★★

정리 3.3, 5.2, 4회 정리 4.2 · 제한 40분

$X_1, \dots, X_n \overset{\text{iid}}{\sim}$ Exp(평균 θ), $n \geq 2$이고 $c > 0$은 알려진 상수다. $\tau(\theta) = P_\theta(X_1 > c) = e^{-c/\theta}$ 의 UMVUE를 구한다. $S = \sum X_i$이다.

(a) S가 완비충분통계량임을 보여라.

(b) $X_1/S \sim$ 베타$(1, n-1)$이고 X_1/S가 S와 독립임을 보여라.

(c) $P(X_1 > c \mid S = s) = (1 - c/s)^{n-1}$ $(s > c)$, 0 $(s \leq c)$임을 보이고 UMVUE를 적어라.

(d) $n = 5$, $\sum x_i = 10$, $c = 1$이면 UMVUE의 값은 얼마인가? MLE $e^{-c/\bar{x}}$와 비교하라.

D2 Poisson에서 "정확히 한 번"의 확률 ★★★

정리 2.1, 3.3, 4.5 · 제한 30분

$X_1, \dots, X_n \overset{\text{iid}}{\sim} \text{Poi}(\lambda)$이고 $\tau(\lambda) = P(X = 1) = \lambda e^{-\lambda}$이다. $T = \sum X_i$이다.

(a) τ의 불편추정량을 하나 찾아라.

(b) (a)에 T로 조건을 걸어 UMVUE $\varphi(T)$를 구하라(답의 변수는 t와 n).

(c) φ가 UMVUE인 근거를 정리 번호로 적어라.

(d) $n = 10$, $T = 7$이면 UMVUE의 값은 얼마인가?

(e) 같은 자료에서 MLE $\tau(\bar{x})$의 값을 구하고 비교하라.

D3 정규분포의 σ와 μ^2 ★★★ 정리 3.3, 4.5, 6회 정리 2.2 · 제한 35분

$X_1, \dots, X_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$, $n \geq 2$이다.

(a) $ES = \sigma\sqrt{\frac{2}{n-1}}\,\frac{\Gamma(n/2)}{\Gamma((n-1)/2)}$ 임을 보여라.

(b) σ의 UMVUE가 $c_n S$, $c_n = \sqrt{\frac{n-1}{2}}\,\frac{\Gamma((n-1)/2)}{\Gamma(n/2)}$ 임을 보여라.

(c) c_5를 닫힌 꼴과 소수로 구하라.

(d) μ^2의 UMVUE를 구하고, $n=5$, $\bar{x}=3$, $s^2=2$일 때 값을 구하라.

(e) (d)의 UMVUE가 음수가 될 수 있는가? 7절의 터무니없는 UMVUE와 비교하라.

D4 균등분포의 분포함수 값 ★★★

정리 2.1, 3.3 · 제한 35분

$X_1, \dots, X_n \overset{\text{iid}}{\sim} \mathrm{U}(0,\theta)$이고 $c > 0$은 알려진 상수다. $\tau(\theta) = P_\theta(X_1 \le c) = \min(c/\theta, 1)$의 UMVUE를 구한다. $M = X_{(n)}$이다.

(a) $M = m$이 주어지면 확률 $\frac{1}{n}$로 $X_1 = m$이고, 나머지 확률로 $X_1 \sim \mathrm{U}(0, m)$임을 설명하라.

(b) $\varphi(m) = E\big[1\{X_1 \le c\} \mid M = m\big]$을 구하라.

(c) $\varphi(M)$이 UMVUE인 근거를 적어라.

(d) $n = 4$, $M = 5$, $c = 2$이면 UMVUE의 값은 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 (a) 분자는 $p^3(1-p)^2$, 분모는 $T \sim \mathrm{b}(5,p)$의 확률이다. (b) 분모는 $T \sim \mathrm{Poi}(3\lambda)$의 확률이다(4회 정리 2.4).

답 (a) $\frac{1}{10}$ (b) $\frac{4}{27}$

A2 힌트 $T=t$가 주어지면 $X_1 \sim \mathrm{b}(t,\frac{1}{n})$이다(노트 연습 1.1). $T=\sum X_i$는 완비충분이다(정리 4.5).

답 (a) $(\frac{3}{4})^6 = \frac{729}{4096} \approx 0.178$ (b) $\frac{1215}{4096} \approx 0.297$

A3 힌트 $T=t$가 주어지면 성공 t개의 자리는 모든 배치가 같은 확률이다. X_1X_2와 $X_1(1-X_2)$에 조건을 건다.

답 (a) $\frac{2}{15}$ (b) $\frac{4}{15}$

A4 힌트 $X_{(n)}$의 밀도는 nt^{n-1}/θ^n이고 완비충분이다. $E\,X_{(n)}^{-1}$을 적분으로 구한다. $n \geq 2$가 필요하다.

답 (a) 3.875 (b) $\frac{15}{62} \approx 0.242$

A5 힌트 UMVUE는 $\frac{k+1}{k}M - 1$이다. 예제 2.3의 조건부 평균의 계산을 떠올린다.

답 (a) 74 (b) 79.5

A6 힌트 $e^{A(\eta)} = \sum_{x\geq 1} e^{\eta x}$이다($h=1$이므로). 기하급수 [T5]. $e^{\eta} = 1-p$로 되돌린다.

답 (a) $A(\eta) = \eta - \log(1-e^{\eta})$ (b) 4 (c) 12

A7 힌트 (a)는 $A''(\eta)$이다. (b)는 9회의 Bernoulli 정보 $\frac{1}{p(1-p)}$이다. 변환 규칙은 $I(p) = I(\eta)\left(\frac{d\eta}{dp}\right)^2$이다.

답 (a) 0.16 (b) 6.25

A8 힌트 $\sum X_i$는 완비충분이고 $X_{(4)}/\sum X_i$는 척도를 바꿔도 같다. Basu로 둘은 독립이다. $X_{(4)} = \frac{X_{(4)}}{\sum X_i} \cdot \sum X_i$의 기댓값을 쓴다.

답 $\frac{25}{48} \approx 0.521$

A9 힌트 예제 5.5(b)와 같다. $X_{(5)}$는 완비충분, $X_{(2)}/X_{(5)}$는 보조다.

답 $\frac{2}{5}$

A10 힌트 $E(1-\frac{1}{n})^{2T}$은 $s=(1-\frac{1}{n})^2$에서의 확률생성함수다. Poisson 한 관측의 정보는 $1/\lambda$이다.

답 (a) $e^{-2}(e^{0.1}-1)\approx 0.01423$ (b) $\frac{e^{-2}}{10}\approx 0.01353$ (c) $10(e^{0.1}-1)\approx 1.0517$

A11 힌트 $g(0)(1-p)^2+2g(1)p(1-p)+g(2)p^2=0$을 두 p에 대해 쓰면 미지수 세 개에 식 두 개다. 두 식을 빼 본다.

답 (a) $g(1)=-\frac{5}{3}$ (b) $g(2)=1$

A12 힌트 $\eta=-\frac{1}{2\theta}$로 두면 $f=x\exp\!\big(\eta x^2+\log(-2\eta)\big)$이다.

답 $\hat\theta=\frac{9}{4}$

A13 힌트 $\mu_2=1+8\varepsilon$, $\mu_4=3(1+80\varepsilon)$, $\nu_1=\sqrt{2/\pi}(1+2\varepsilon)$. 효율은 $\frac{1}{4}(\mu_4/\mu_2^2-1)$을 μ_2/ν_1^2-1로 나눈 것이다.

답 ≈ 1.751

A14 힌트 $T\sim\text{Bernoulli}(q(\lambda))$, $q=e^{-\lambda}$이다. Bernoulli 정보의 변환 규칙 $I_T=\frac{q'(\lambda)^2}{q(1-q)}$. $I_X=1/\lambda$.

답 (a) $I_T(\lambda)=\frac{1}{e^{\lambda}-1}$ (b) $\frac{2}{e^{2}-1}\approx 0.313$

B1 힌트 근거는 짧게 쓴다. 예: "1회 정리 5.3", "정의 1.1". ($\Leftarrow$)의 첫 단계는 $\{T=t\}$를 서로소인 사건 $\{\mathbf{X}=\mathbf{y}\}$로 쪼갠 것이다.

B2 힌트 충분성은 한 단계에서만 쓰인다. 어느 단계인지 먼저 찾는다.

B3 힌트 완비성은 "모든 θ에서 평균이 0인 T의 함수는 0"이다. 그 함수가 무엇인지 찾는다.

B4 힌트 B1을 다시 보지 말고 쓴다. 약분의 조건 $g(t;\theta)>0$이 어디서 오는지 밝힌다.

B5 힌트 등호 조건에서는 $E\,\mathrm{Var}(W\mid T)=E(W-\varphi(T))^2$을 반복기댓값으로 보인다.

B6 힌트 단계는 일곱이다. 다 쓴 뒤 $P_\theta(T=t)=0$인 경우를 따로 처리했는지 확인한다.

B7 힌트 다 쓴 뒤, 보조성·충분성·완비성이 각각 어느 단계에서 쓰였는지 표시한다. 셋 가운데 하나라도 빠지면 어느 단계가 무너지는가?

C1 힌트 X_1과 X_2는 독립이므로 $X_2\mid X_1=x_1\sim\mathcal{N}(\theta,1)$이다.

C2 힌트 (a) 로그 가능도비는 $-\frac{1}{2\theta^2}(\sum x_i^2-\sum y_i^2)+\frac{1}{\theta}(\sum x_i-\sum y_i)$이다. $1/\theta$와 $1/\theta^2$은 일차종

속이 아니다. (b) $E\bar{X}^2 = \theta^2 + \theta^2/n$.

C3 힌트 (a) $f = e^{n\theta - \sum x_i} 1\{x_{(1)} > \theta\}$. (c) $\int_\theta^\infty g(t)e^{-nt}dt = 0$이 모든 θ에서 성립하면 θ로 미분한다(미적분의 기본정리).

답 (d) 2.1

C4 힌트 $E_\theta \varphi(T) = E_\theta W$이므로 두 추정량의 치우침은 같다. 평균제곱오차 = 분산 + 치우침^2.

C5 힌트 (a) R은 $X_i - \mu$들만의 함수다. (b) $(\bar{X}, S)$는 완비충분(정리 4.5), R/S는 위치와 척도를 바꿔도 같다. (c) $RS = (R/S)S^2$에 (b)를 쓴다.

C6 힌트 (a) 인수분해 $f = g(T;\theta)h = g(r(U);\theta)h$. (c) $(X_1, \bar{X})$는 결합정규이고 $\mathrm{Var}(X_1 \mid \bar{X}) = \sigma^2 \frac{n-1}{n}$이다(5회 정리 C).

C7 힌트 (a) $\frac{d\mu}{d\eta} = A''(\eta) = \mathrm{Var}\, T$. (b) 변환 규칙 $I(\mu) = I(\eta)\left(\frac{d\eta}{d\mu}\right)^2$.

C8 힌트 (a) $W_a = W + aU$도 불편이다. $\mathrm{Var}\, W_a$는 a의 이차식이다. $\mathrm{Cov} \neq 0$인 θ_0에서 $a = -\mathrm{Cov}/\mathrm{Var}\, U$로 두어 귀류 [T2]. (b) 다른 불편추정량 W'에 $U = W' - W$.

C9 힌트 연습 1.6의 (b)에서 등호이면 $E\,\mathrm{Var}(s_{\mathbf{X}} \mid T) = 0$이다. 그러면 $s_{\mathbf{X}}(\mathbf{x};\theta) = s_T(T(\mathbf{x});\theta)$이고, 이것은 $\partial_\theta \log P_\theta(\mathbf{X} = \mathbf{x} \mid T = T(\mathbf{x})) = 0$이라는 뜻이다.

D1 힌트 (b) X_1은 감마$(1,\theta)$, $\sum_{i\geq 2} X_i$는 감마$(n-1,\theta)$이고 독립이다. 4회의 베타–감마 관계. 독립은 Basu로도 나온다(X_1/S는 척도에 불변). (c) 베타$(1,n-1)$의 꼬리는 $P(B > u) = (1-u)^{n-1}$.

답 (c) $(1 - c/S)^{n-1}$ $(S > c)$ (d) $0.9^4 = 0.6561$

D2 힌트 $W = 1\{X_1 = 1\}$. $T = t$가 주어지면 $X_1 \sim \mathrm{b}(t, \frac{1}{n})$이다.

답 (b) $\frac{T}{n}(1 - \frac{1}{n})^{T-1}$ (d) $0.7 \cdot 0.9^6 \approx 0.3720$ (e) $0.7e^{-0.7} \approx 0.3476$

D3 힌트 (a) $(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$이므로 $E\sqrt{V}$, $V \sim \chi^2(k)$를 감마 적분으로 구한다. (d) $E\bar{X}^2 = \mu^2 + \sigma^2/n$.

답 (c) $c_5 = \frac{4\sqrt{2}}{3\sqrt{\pi}} \approx 1.0638$ (d) 8.6

D4 힌트 (a) 대칭으로 최댓값이 X_1일 확률은 $\frac{1}{n}$이고, 최댓값이 아닌 관측은 $(0,m)$에서 균등하다(8회의 순서통계량). (b) $m \leq c$와 $m > c$를 나눈다.

답 (d) $\frac{3}{10}$