수리통계학 × 통계학의 역사

16회 워크북

베이즈의 귀환
손으로 푸는 34문항

강의노트 16회를 읽은 뒤 손으로 푼다. 사후분포는 사전밀도와 가능도를 곱해 핵을 알아보는 계산이고, MCMC와 James–Stein의 증명은 각각 한 줄의 대칭과 한 번의 부분적분이다. 그 계산과 그 한 줄을 빈 종이에 다시 쓰는 것이 이 워크북의 일이다. 문제 아래의 줄 친 공간에 직접 쓴다.

부	하는 일	문항	예상 시간
A 계산 드릴	사후분포, Bayes 추정량, Bayes 인수, 전이핵, 축소를 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 40분
D 퀄형	여러 회차를 잇는 종합 문제. 제한 시간 안에	4	2시간 25분
	합계	34	약 10시간

노트 `16_bayes_mcmc_notes.pdf` — 문항의 "정리 5.2" 같은 번호는 노트의 번호다

도구 `00_toolkit_notes.pdf` — [T2] 귀류·반례, [T4] 부등식, [T5] 극한, [T6] 극한과 적분의 교환, [T7] 변수변환

검산 노트북 `16_bayes_mcmc.ipynb`의 "워크북 검산" 절: `check("A1a", "3/5")`

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, θ는 `theta`로, e는 `e`로 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. 사후분포는 먼저 핵을 알아보고, 적분은 마지막에 한다.

A1 베타–이항의 사후분포 ★

정리 1.2 · 8분

$p \sim \mathrm{Beta}(2,3)$이고 $S \mid p \sim \mathrm{b}(10,p)$에서 $s=7$을 관측했다. (a) 사후평균, (b) 사후 최빈값(MAP), (c) 사후분산을 구하라. 사후평균이 사전평균과 표본비율 사이에 있음을 확인하라.

A2 감마–Poisson과 예측 ★

정리 1.2 · 8분

$\lambda \sim \Gamma(2,\ \text{율}\ 1)$이고 $X_1,\dots,X_5 \mid \lambda \overset{\text{iid}}{\sim} \mathrm{Poi}(\lambda)$에서 $3,1,4,0,2$를 관측했다. (a) λ의 사후평균, (b) 다음 관측이 0일 사후예측확률 $P(X_6=0 \mid \mathbf{x})$를 구하라. MLE를 넣은 $e^{-\hat\lambda}$와 비교하라.

A3 정규–정규의 축소 ★

정리 1.2(c) · 8분

$\mu \sim \mathcal{N}(0, 4)$이고 $X_1, \dots, X_4 \mid \mu \overset{\text{iid}}{\sim} \mathcal{N}(\mu, 1)$에서 $\bar{x} = 2.5$를 관측했다. (a) 사후평균, (b) 사후분산을 구하라. $\bar{x}$에 주는 무게는 얼마인가?

A4 손실마다 다른 Bayes 추정량 ★★

정리 2.1 · 10분

사후밀도가 $\pi(\theta \mid \mathbf{x}) = 2\theta\,(0 < \theta < 1)$이다. 다음 손실의 Bayes 추정량을 구하라. (a) 제곱손실, (b) 절대손실, (c) 0–1 손실에서 $\epsilon \to 0$의 극한(MAP), (d) 가중 제곱손실 $L = (\theta - a)^2/\theta$.

A5 상수 위험의 Bayes 추정량 ★★

예제 2.2, 정리 2.3 · 10분

$S \sim \mathrm{b}(16, p)$이고 Beta(2, 2) 사전분포의 사후평균 $\delta = \frac{S+2}{20}$을 쓴다. (a) δ의 위험이 p에 무관함을 보이고 그 값을 구하라. (b) MLE $S/16$의 위험의 최댓값을 구하라. (c) MLE의 위험이 δ보다 작은 p의 범위는 $p < p_0$ 또는 $p > 1 - p_0$이다. p_0을 구하라.

A6 지수분포의 Jeffreys 사전분포 ★★

정의 2.5, 정리 2.6 · [T7] · 10분

$f(x \mid \theta) = \theta e^{-\theta x}$ $(x > 0)$이다. (a) Jeffreys 사전분포가 θ^k에 비례할 때 k를 구하라. (b) 관측 5개의 합이 8이면 이 사전분포에서 θ의 사후평균은? (c) 평균 $\mu = 1/\theta$로 모수를 바꿔 Jeffreys 사전분포를 μ에서 직접 만들면 μ^k의 k는? 정리 2.6의 결론과 맞는지 변수변환으로 확인하라.

A7 이항의 Bayes 인수 ★

정의 3.1, 명제 3.2 · 8분

10번 가운데 8번 성공했다. $H_0 : p = \frac{1}{2}$, $H_1 : p \sim \mathrm{U}(0,1)$이다. (a) B_{10}, (b) 증거의 무게(deciban), (c) 두 가설에 사전확률을 반씩 줄 때 $P(H_0 \mid \mathbf{x})$를 구하라. 정확한 양측 유의확률 $2P(S \geq 8)$과 비교하라.

A8 세 단서 ★

명제 3.2 · 6분

서로 독립인 세 단서의 무게가 5, 10, 20 deciban이다. 사전 오즈가 $1 : 100$이다. (a) 합친 무게를 구하라. (b) 사후확률을 구하라.

A9 세 상태 연쇄의 정상분포 ★

정의 4.1, 4.2 · 6분

상태 $\{1,2,3\}$의 전이행렬이 다음과 같다. (a) π_1, (b) π_2를 구하고, 상세 균형이 성립함을 확인하라.

$$P = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 1/2 & 1/2 \end{pmatrix}$$

A10 받아들임 확률 손 계산 ★

알고리즘 4.4 · 8분

목표가 $\pi(x) \propto x^2 e^{-x}$ $(x > 0)$이다. (a) 대칭 제안으로 $x = 2$에서 $y = 3$이 제안될 때, (b) $y = 1$이 제안될 때의 받아들임 확률을 구하라. (c) 독립 제안 $g(y) = e^{-y}$를 쓸 때 $x = 2$에서 $y = 1$이 제안되면 받아들임 확률은?

A11 2×2 표의 Gibbs ★★

알고리즘 4.6, 정리 4.7 · 10분

$(X, Y) \in \{0,1\}^2$의 결합확률이 $p(0,0) = 0.4$, $p(0,1) = 0.1$, $p(1,0) = 0.2$, $p(1,1) = 0.3$이다. 한 바퀴는 $Y \sim p(y \mid X)$, 이어서 $X' \sim p(x \mid Y)$다. (a) $P(X' = 1 \mid X = 0)$, (b) $P(X' = 1 \mid X = 1)$, (c) X열의 정상분포에서 $P(X = 1)$을 구하라. X의 주변분포와 비교하라.

A12 Stein의 보조정리로 넷째 적률 ★

보조정리 5.1 · 8분

$Z \sim \mathcal{N}(\theta, 1)$이다. 보조정리 5.1을 $g(z) = z^3$에 써서 EZ^4을 θ의 식으로 나타내라. $EZ^2 = \theta^2 + 1$, $EZ^3 = \theta^3 + 3\theta$는 써도 된다(노트 연습 5.2).

A13 James–Stein 손 계산 ★

정리 5.2 · 8분

$p = 5$이고 $\mathbf{X} = (1, -2, 0.5, 1.5, -1)$이다. (a) 축소 인자 $1 - \frac{p-2}{\|\mathbf{X}\|^2}$, (b) δ_{JS}의 둘째 성분을 구하라. (c) $\boldsymbol{\theta} = \mathbf{0}$에서 δ_{JS}의 위험은?

A14 경험적 Bayes의 숫자 ★

정리 5.4 · 8분

$\theta_i \overset{\text{iid}}{\sim} \mathcal{N}(0, \tau^2)$, $X_i \mid \theta_i \sim \mathcal{N}(\theta_i, 1)$, $p = 10$이고 $\|\mathbf{X}\|^2 = 40$을 관측했다. (a) $B = \frac{1}{1+\tau^2}$의 불편추정값 $\frac{p-2}{\|\mathbf{X}\|^2}$, (b) 그 값에 해당하는 τ^2, (c) 참값이 $\tau^2 = 4$일 때 δ_{JS}의 Bayes 위험을 구하라. Bayes 추정량의 Bayes 위험과 비교하라.

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 결과, 도구 번호를 쓴다.

B1 James–Stein, 근거 채우기 ★★

① · 정리 5.2 · 20분

$\mathbf{X} \sim \mathcal{N}_p(\theta, I)$, $p \geq 3$이면 $\delta_{\mathrm{JS}} = \left(1 - \frac{p-2}{\|\mathbf{X}\|^2}\right)\mathbf{X}$의 위험은 $p - (p-2)^2 E[1/\|\mathbf{X}\|^2]$이다. 각 단계의 근거를 써라.

① $\mathbf{g}(\mathbf{x}) = (p-2)\mathbf{x}/\|\mathbf{x}\|^2$로 두면 $\delta_{\mathrm{JS}} = \mathbf{X} - \mathbf{g}(\mathbf{X})$이다. 정의

② $\|\delta_{\mathrm{JS}} - \theta\|^2 = \|\mathbf{X} - \theta\|^2 - 2(\mathbf{X} - \theta)'\mathbf{g}(\mathbf{X}) + \|\mathbf{g}(\mathbf{X})\|^2$이다. 왜? _____

③ $\nabla \cdot \mathbf{g}(\mathbf{x}) = (p-2)^2/\|\mathbf{x}\|^2$이다. 왜? _____

④ $E[(\mathbf{X} - \theta)'\mathbf{g}(\mathbf{X})] = (p-2)^2 E[1/\|\mathbf{X}\|^2]$이다. 왜? _____

⑤ $\|\mathbf{g}(\mathbf{X})\|^2 = (p-2)^2/\|\mathbf{X}\|^2$이다. 왜? _____

⑥ 기댓값을 더하면 $p - 2(p-2)^2 E[\cdot] + (p-2)^2 E[\cdot] = p - (p-2)^2 E[1/\|\mathbf{X}\|^2]$이다. 왜? _____

⑦ $p \geq 3$이면 이 기댓값은 양수이고 유한하므로 위험은 p보다 작다. 왜? _____

B2 Metropolis–Hastings, 근거 채우기 ★

① · 정리 4.5 · 15분

알고리즘 4.4의 전이핵이 π에 대해 상세 균형을 만족함의 증명이다. 근거를 채워라.

① $x \neq y$이면 $k(x,y) = q(x,y)\alpha(x,y)$이다. 왜? _____

② $\pi(x)k(x,y) = \min\{\pi(x)q(x,y),\ \pi(y)q(y,x)\}$이다. 왜? _____

③ 따라서 $\pi(x)k(x,y) = \pi(y)k(y,x)$이다. 왜? _____

④ $x = y$의 제자리 덩어리는 상세 균형을 어기지 않는다. 왜? _____

⑤ π는 정상분포다. 왜? _____

B3 Stein의 보조정리, 근거 채우기 ★★

① · 보조정리 5.1 · 20분

$Z \sim \mathcal{N}(0,1)$, g 절대연속, $E|g'(Z)| < \infty$이면 $E[Zg(Z)] = E[g'(Z)]$임의 증명이다. 근거를 채워라.

① $z > 0$에서 $\varphi(z) = \int_z^\infty u\varphi(u)\,du$이다. 왜? _____

② $\int_0^\infty g'(z)\varphi(z)\,dz = \int_0^\infty u\varphi(u)\int_0^u g'(z)\,dz\,du$이다. 왜? _____

③ $= \int_0^\infty u\varphi(u)\big(g(u) - g(0)\big)du$이다. 왜? _____

④ 음의 쪽도 같아 $E[g'(Z)] = E[Z(g(Z) - g(0))]$이다. 왜? _____

⑤ $= E[Zg(Z)]$이다. 왜? _____

B4 James–Stein, 단계의 이름만 ★★

② · 정리 5.2 · 25분

정리 5.2를 증명하라. 단계의 이름만 주어진다. 각 단계 아래에 식과 근거를 써라.

① **δ_{JS}를 $\mathbf{X}$에서 함수 $\mathbf{g}$를 뺀 꼴로 쓴다**

② **손실을 세 항으로 전개한다**

③ **g의 발산을 계산한다**

④ **교차항을 Stein의 보조정리로 바꾼다**

⑤ **기댓값을 모아 위험을 쓴다**

⑥ **$p \geq 3$에서 위험이 p보다 작음을 결론짓는다**

B5 Metropolis–Hastings, 단계의 이름만 ★★

② · 정리 4.5 · 25분

정리 4.5를 증명하라. 단계의 이름만 주어진다.

① **옮기는 확률을 제안과 받아들임으로 쓴다**

② **$\pi(x)$를 받아들임 확률의 min 안으로 넣는다**

③ **대칭에서 상세 균형을 읽는다**

④ **제자리 덩어리를 따로 본다**

⑤ **정상분포를 결론짓는다**

⑥ $\pi(x)q(x,y)=0$**인 경우를 따진다**

B6 James–Stein, 전략만 ★★★

③ · 정리 5.2 · 30분

정리 5.2를 증명하라. 원점의 특이점이 보조정리 5.1의 조건을 깨지 않는 이유도 밝혀라.

전략 $\delta_{\mathrm{JS}}=\mathbf{X}-\mathbf{g}(\mathbf{X})$로 쓰고 손실을 전개한다. 모르는 θ가 들어 있는 교차항을 Stein의 보조정리로 θ가 없는 기댓값으로 바꾼다.

B7 Stein의 보조정리, 전략만 ★★★

③ · 보조정리 5.1 · 30분

보조정리 5.1의 1차원 판을 평균이 θ인 경우까지 증명하라. Fubini를 쓰는 곳에서 그 조건을 확인하라.

전략 부분적분을 Fubini로 한다. 정규밀도를 $\varphi(z) = \int_z^\infty u\varphi(u)\,du$로 쓰면 $E[g'(Z)]$는 이중적분이 되고, 적분 순서를 바꾸면 g'의 적분이 g로 돌아온다.

C

증명 문제

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

C1 균등 사전분포는 모수화에 달려 있다 ★

정리 2.6 · [T7] · 15분

$X \sim \text{Bernoulli}(p)$이다.

(a) $p \sim \text{U}(0,1)$이면 $\phi = p^2$의 밀도가 $\frac{1}{2\sqrt{\phi}}$ $(0 < \phi < 1)$이고 균등이 아님을 보여라.

(b) $\phi = p^2$을 모수로 써서 Fisher 정보 $I_\phi(\phi)$를 직접 계산하고 Jeffreys 사전분포 $\sqrt{I_\phi}$를 구하라.

(c) (b)가 $\text{Beta}(\frac{1}{2}, \frac{1}{2})$를 $\phi = p^2$으로 변수변환한 밀도와 상수배를 빼고 같음을 확인하라.

C2 상수 추정량은 허용적이다 ★★

[T2] · 정리 2.3 · 15분

$X \sim \mathcal{N}(\theta, 1)$, 제곱손실이다. 상수 추정량 $\delta_0(X) \equiv 0$은 허용적이지만 minimax가 아님을 보여라. 허용성이 "좋은 추정량"의 충분조건이 아님을 이 예로 설명하라.

C3 Bayes 위험의 극한과 minimax ★★

[T5] · 정리 2.3 · 20분

(a) 사전분포의 열 π_k와 그 Bayes 추정량 δ_k에 대해 $r(\pi_k, \delta_k) \to c$이고 어떤 추정량 δ가 $\sup_\theta R(\theta, \delta) = c$이면 δ는 minimax임을 보여라.

(b) $X_1, \dots, X_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, 1)$, 사전분포 $\mathcal{N}(0, \tau^2)$의 Bayes 위험 r_τ를 τ와 n의 식으로 구하라.

(c) (a)와 (b)로 $\bar{X}$가 minimax임을 보여라. 정리 2.3(b)를 바로 쓸 수 없는 이유는?

C4 Barker의 받아들임 ★★

[T4] · 정리 4.5 · 20분

$r(x,y)=\frac{\pi(y)q(y,x)}{\pi(x)q(x,y)}$로 두고, MH의 $\alpha=\min\{1,r\}$ 대신 Barker(1965)의 $\alpha_B=\frac{r}{1+r}$을 쓴다.

(a) Barker의 전이핵도 π에 대해 상세 균형을 만족함을 보여라.

(b) 모든 $r>0$에서 $\alpha_B\le\min\{1,r\}$임을 보여라. 같은 제안이면 MH가 덜 머무는 이유다.

(c) 대칭 제안에서 $\pi(y)/\pi(x)=3$이면 Barker의 받아들임 확률은?

C5 두 블록 Gibbs의 주변 연쇄는 가역이다 ★★

정리 4.7 · 20분

결합분포 $\pi(x,y)$에서 두 블록 Gibbs를 돌린다. 한 바퀴는 $Y\sim\pi(y\mid X)$, $X'\sim\pi(x\mid Y)$다.

(a) X열의 전이핵이 $K(x,x')=\int\pi(y\mid x)\pi(x'\mid y)\,dy$임을 보여라.

(b) $\pi_X(x)K(x,x')=\int\frac{\pi(x,y)\pi(x',y)}{\pi_Y(y)}\,dy$이고, 따라서 X열은 π_X에 대해 가역임을 보여라.

(c) 체계적 순서의 한 바퀴는 일반적으로 가역이 아니다(노트 4절). (b)와 어긋나지 않는 이유를 설명하라.

C6 Poisson의 Stein 항등식 ★

보조정리 5.1 · 15분

$X \sim \mathrm{Poi}(\lambda)$이다.

(a) $E|Xg(X-1)| < \infty$이면 $E[\lambda g(X)] = E[Xg(X-1)]$임을 보여라. (정규의 $E[(Z-\theta)g(Z)] = E[g'(Z)]$와 짝을 이루는 이산판이다.)

(b) $g(x) = x + 1$을 넣어 EX^2을 λ의 식으로 구하라. (λ는 `lam`으로 쓴다.)

C7 임의의 점을 향한 축소 ★★

정리 5.2 · 20분

$\boldsymbol{\nu} \in \mathbb{R}^p$를 미리 정한 점이라 하자. $p \geq 3$이면

$$\delta_{\boldsymbol{\nu}}(\mathbf{X}) = \boldsymbol{\nu} + \Big(1 - \frac{p-2}{\|\mathbf{X}-\boldsymbol{\nu}\|^2}\Big)(\mathbf{X}-\boldsymbol{\nu})$$

의 위험이 $p - (p-2)^2 E\big[1/\|\mathbf{X}-\boldsymbol{\nu}\|^2\big] < p$임을 보여라. 위험이 가장 작은 θ는 어디인가?

C8 축소 상수의 범위 ★★

정리 5.2 · 15분

$c > 0$에 대해 $\delta_c(\mathbf{X}) = \big(1 - \frac{c}{\|\mathbf{X}\|^2}\big)\mathbf{X}$, $p \geq 3$이다.

(a) $R(\theta, \delta_c) = p - \big[2c(p-2) - c^2\big] E[1/\|\mathbf{X}\|^2]$임을 보여라.

(b) δ_c가 $\mathbf{X}$를 지배할 필요충분조건이 $0 < c < 2(p-2)$임을 보여라.

(c) 모든 θ에서 위험이 가장 작은 c를 p의 식으로 구하라.

C9 두 점 사전분포의 Tweedie ★★

정리 8.5 · 20분

θ가 ± 1을 반씩 갖고 $X \mid \theta \sim \mathcal{N}(\theta, 1)$이다. $E(\theta \mid X = x)$를 (a) 베이즈 정리로 직접, (b) Tweedie의 공식으로 구해 같음을 보여라. (c) 이 사후평균이 x가 클 때 1로 가는 속도를 말하라. (x에 대한 식으로 답한다.)

D

퀄형

여러 회차의 도구를 잇는 종합 문제다. 머리의 제한 시간을 재고, 노트를 덮고 푼다. 모형과 손실을 식으로 적는 것부터 답안의 일부다.

D1 $e^{-\lambda}$의 Bayes 추정과 UMVUE ★★★

정리 1.2, 2.1 · 11회 예제 2.2 · [T8] · 제한 35분

$X_1, \dots, X_n \mid \lambda \overset{\text{iid}}{\sim} \text{Poi}(\lambda)$, $\lambda \sim \Gamma(a,$ 율 $b)$이고 $T = \sum X_i$이다. 추정할 양은 $\tau = e^{-\lambda} = P(X = 0)$이다.

(a) 제곱손실에서 τ의 Bayes 추정량이 $\left(\frac{b+n}{b+n+1}\right)^{a+T}$ 임을 보여라.

(b) $a = b = 1$, $n = 10$, $T = 12$이면 그 값은?

(c) 같은 자료에서 11회 예제 2.2의 UMVUE $(1 - \frac{1}{n})^T$의 값은?

(d) $n \to \infty$이면 두 추정량과 $e^{-T/n}$의 차이가 모두 $O_p(1/n)$임을 보여라.

D2 가중 손실에서 MLE는 minimax다 ★★★ 정리 2.1(c), 2.3 · 제한 35분

$S \sim \mathrm{b}(n,p)$이고 손실은 $L(p,a)=\dfrac{(p-a)^2}{p(1-p)}$, 사전분포는 $\mathrm{U}(0,1)$이다.

(a) $0<s<n$이면 Bayes 추정량이 s/n임을 보여라.

(b) $s=0$과 $s=n$에서도 Bayes 추정량이 s/n임을 보여라. (사후 기대손실이 $a \neq s/n$에서 무한함을 보인다.)

(c) S/n의 위험이 p에 무관한 상수임을 보이고 그 값을 구하라.

(d) S/n이 이 손실에서 minimax임을 결론짓고, 제곱손실에서는 minimax가 아니었음(A5)과 비교하라.

D3 기하분포를 목표로 한 Metropolis ★★★

알고리즘 4.4, 정리 4.5 · 제한 35분

목표는 $\pi(k) = (1-r)r^k$ $(k = 0, 1, 2, \dots,\ 0 < r < 1)$이고, 현재 상태 k에서 $k-1$과 $k+1$을 반씩 제안한다(-1은 목표 밖이므로 늘 거절된다).

(a) $\alpha(k, k+1)$과 $\alpha(k+1, k)$를 구하라.

(b) 이 연쇄가 π에 대해 상세 균형을 만족함을 직접 확인하라.

(c) $P(X_{t+1} = 0 \mid X_t = 0)$을 r의 식으로 구하라.

(d) $k \geq 1$에서 $P(X_{t+1} = k \mid X_t = k)$를 구하라.

(e) 0에서 출발해 처음 1로 옮겨 가기까지 걸리는 걸음 수의 기댓값을 구하라.

D4 비정칙 사전분포와 t 사후분포 ★★★

6회 정리 3.2, 5.2 · [T7] · 제한 40분

$X_1, \dots, X_n \mid \mu, \sigma^2 \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$, 사전분포는 $\pi(\mu, \sigma^2) \propto 1/\sigma^2$ 이다. $\bar{x}$, s^2 은 표본평균과 표본분산이다.

(a) σ^2 을 적분해 μ 의 주변 사후밀도가 $\left[(n-1)s^2 + n(\mu - \bar{x})^2\right]^{-n/2}$ 에 비례함을 보여라.

(b) $\dfrac{\mu - \bar{x}}{s/\sqrt{n}} \,\Big|\, \mathbf{x} \sim t_{n-1}$ 임을 보여라.

(c) $n = 10$, $\bar{x} = 5$, $s = 2$ 이면 95% 등꼬리 신용구간의 위 끝은? ($t_{9,0.975} = 2.2622$)

(d) 같은 자료에서 $P(\mu > 6 \mid \mathbf{x})$ 를 구하라.

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 사후분포는 $\text{Beta}(2+7,\ 3+3)$이다. $\text{Beta}(a,b)$의 평균은 $\frac{a}{a+b}$, 최빈값은 $\frac{a-1}{a+b-2}$, 분산은 $\frac{ab}{(a+b)^2(a+b+1)}$이다.

답 (a) $\frac{3}{5}$ (b) $\frac{8}{13}$ (c) $\frac{3}{200}$

A2 힌트 사후분포는 $\Gamma(2+10,\ \text{율}\ 1+5)$이다. (b)는 $E(e^{-\lambda}\mid \mathbf{x})$이고, 감마 밀도의 적분으로 계산한다.

답 (a) 2 (b) $(6/7)^{12}\approx 0.157$

A3 힌트 정밀도를 더한다. 자료의 정밀도는 $n/\sigma^2=4$, 사전분포의 정밀도는 $1/\tau^2=\frac{1}{4}$이다.

답 (a) $\frac{40}{17}\approx 2.353$ (b) $\frac{4}{17}$

A4 힌트 (b) 사후분포함수 θ^2이 $\frac{1}{2}$이 되는 점. (d) 정리 2.1(c)에서 $w(\theta)=1/\theta$이다.

답 (a) $\frac{2}{3}$ (b) $1/\sqrt{2}\approx 0.707$ (c) 1 (d) $\frac{1}{2}$

A5 힌트 예제 2.2의 위험 공식에 $n=16$, $a=b=2=\sqrt{n}/2$를 넣는다. (c)는 $p(1-p)/16$과 (a)의 값을 같게 놓는다.

답 (a) $\frac{1}{100}$ (b) $\frac{1}{64}$ (c) $p<0.2$ 또는 $p>0.8$

A6 힌트 점수는 $1/\theta - x$이고 그 분산이 정보다. (b)의 사후밀도는 $\theta^{-1}\cdot\theta^{5}e^{-8\theta}$에 비례한다.

답 (a) $k=-1$ (b) $\frac{5}{8}$ (c) $k=-1$

A7 힌트 $m_1=\int_0^1 \binom{10}{8}p^8(1-p)^2\,dp=\frac{1}{11}$ (균등 사전분포의 주변 가능도는 모든 s에서 $\frac{1}{n+1}$이다).

답 (a) $\frac{1024}{495}\approx 2.07$ (b) 3.16 (c) $\frac{495}{1519}\approx 0.326$

A8 힌트 사전 오즈 $1:100$은 -20 deciban이다. 무게는 더해진다(명제 3.2(b)).

답 (a) 35 (b) 0.969

A9 힌트 이웃한 상태 사이에서만 움직이므로 상세 균형 $\pi_1P_{12}=\pi_2P_{21}$, $\pi_2P_{23}=\pi_3P_{32}$부터 푼다.

답 (a) $\frac{1}{4}$ (b) $\frac{1}{2}$

A10 힌트 정규화 상수는 비에서 지워진다. (c)는 노트 연습 4.3의 $w = \pi/g$를 쓴다.

답 (a) $\frac{9}{4e} \approx 0.828$ (b) $\frac{e}{4} \approx 0.680$ (c) $\frac{1}{4}$

A11 힌트 $P(X' = 1 \mid X = x) = \sum_y p(y \mid x)\, p(X = 1 \mid y)$. $p(X = 1 \mid Y = 0) = 0.2/0.6$이다.

답 (a) $\frac{5}{12}$ (b) $\frac{7}{12}$ (c) $\frac{1}{2}$

A12 힌트 $E[(Z - \theta)Z^3] = E[3Z^2]$이므로 $EZ^4 = \theta EZ^3 + 3EZ^2$이다.

답 $\theta^4 + 6\theta^2 + 3$

A13 힌트 (c) $\theta = \mathbf{0}$이면 $\|\mathbf{X}\|^2 \sim \chi^2_5$이고 $E[1/\chi^2_p] = 1/(p-2)$이다.

답 (a) $\frac{11}{17}$ (b) $-\frac{22}{17}$ (c) 2

A14 힌트 (c) 정리 5.4(c)의 $p(1 - B) + 2B$.

답 (a) $\hat{B} = \frac{1}{5}$ (b) $\hat{\tau}^2 = 4$ (c) 8.4

B1 힌트 ③은 몫의 미분 $\partial_i (x_i/\|\mathbf{x}\|^2) = (\|\mathbf{x}\|^2 - 2x_i^2)/\|\mathbf{x}\|^4$을 i에 대해 더한 것이다. 교차항에서 모르는 θ를 없애는 것이 Stein의 보조정리다.

B2 힌트 ②는 α의 정의에 $\pi(x)q(x,y)$를 곱해 min 안으로 넣은 것이다.

B3 힌트 ②에서 적분 영역 $\{0 < z < u\}$를 z 먼저에서 u 먼저로 바꿔 읽는다.

B4 힌트 B1에서 채운 근거를 다시 보지 말고 쓴다. 넷째 단계에서 왜 θ가 사라지는지 한 문장으로 적는다.

B5 힌트 마지막 단계: $\pi(y) = 0$인 제안은 $\alpha = 0$으로 거절되므로 연쇄는 $\pi > 0$인 곳에만 머문다.

B6 힌트 특이점: $|\partial_i g_i(\mathbf{x})| \le 3(p-2)/\|\mathbf{x}\|^2$이고, $\|\mathbf{x}\|^{-2}$을 원점 근처 공에서 적분하면 극좌표로 $\int_0^1 r^{p-3} dr$이다 [T7].

B7 힌트 평균 θ는 $h(u) = g(u + \theta)$로 옮긴다. 조건 확인은 $\iint |g'(z)| u\varphi(u) \mathbf{1}\{0 < z < u\}\, dz\, du$가 유한한지 보는 것이다.

C1 힌트 (b) $p = \sqrt{\phi}$이므로 $dp/d\phi = \frac{1}{2\sqrt{\phi}}$이고 $I_\phi = I_p\,(dp/d\phi)^2$이다. 직접 계산하려면 $\log f = x \log\sqrt{\phi} + (1-x)\log(1-\sqrt{\phi})$를 ϕ로 미분한다.

C2 힌트 δ가 δ_0을 지배하면 $R(0,\delta) \le R(0,\delta_0) = 0$이다. 위험이 0이면 $\delta(X) = 0$ (확률 1)이고, $\mathcal{N}(\theta, 1)$들은 같은 영집합을 갖는다.

C3 힌트 (a) 임의의 δ'에 대해 $\sup_\theta R(\theta,\delta') \ge r(\pi_k,\delta') \ge r(\pi_k,\delta_k)$이고 $k\to\infty$로 보낸다. (b) Bayes 위험은 사후분산(정리 2.1(a))이다.

답 (b) $\dfrac{\tau^2}{n\tau^2+1}$

C4 힌트 (a) $\pi(x)q(x,y)\cdot\dfrac{r}{1+r} = \dfrac{\pi(x)q(x,y)\pi(y)q(y,x)}{\pi(x)q(x,y)+\pi(y)q(y,x)}$는 x, y에 대칭이다.

답 (c) $\dfrac{3}{4}$

C5 힌트 (b) $\pi_X(x)\pi(y\mid x)=\pi(x,y)$이고 $\pi(x'\mid y)=\pi(x',y)/\pi_Y(y)$이다. 오른쪽 적분은 x와 x'에 대칭이다.

C6 힌트 $E[Xg(X-1)]=\sum_{k\ge1} k\,g(k-1)e^{-\lambda}\lambda^k/k!$에서 $j=k-1$로 첨자를 옮긴다.

답 (b) $\lambda^2+\lambda$

C7 힌트 $\mathbf{Y}=\mathbf{X}-\boldsymbol{\nu}\sim\mathcal{N}_p(\boldsymbol{\theta}-\boldsymbol{\nu},I)$로 옮기면 정리 5.2 그대로다.

C8 힌트 정리 5.2의 증명에서 $p-2$를 c로 바꾼다. 교차항은 $c(p-2)E\|\mathbf{X}\|^{-2}$, 마지막 항은 $c^2E\|\mathbf{X}\|^{-2}$이다.

답 (c) $c=p-2$

C9 힌트 $m(x)=\frac12\varphi(x-1)+\frac12\varphi(x+1)$이고 $\varphi(x\mp1)\propto e^{\pm x}e^{-(x^2+1)/2}$이다.

답 $\tanh x$

D1 힌트 (a) 사후분포는 $\Gamma(a+T,\ \text{율}\ b+n)$이고 $E(e^{-\lambda}\mid\mathbf{x})$는 그 mgf를 -1에서 계산한 것이다. (d) $\log(1-u)=-u+O(u^2)$와 $T/n\to\lambda$.

답 (a) $\left(\dfrac{b+n}{b+n+1}\right)^{a+T}$ (b) $(11/12)^{13}\approx0.323$ (c) $0.9^{12}\approx0.282$

D2 힌트 (a) 정리 2.1(c)에서 $w(p)=\dfrac{1}{p(1-p)}$이고 사후분포는 $\text{Beta}(s+1,n-s+1)$이다. w를 곱하면 $\text{Beta}(s,n-s)$의 핵이 된다.

답 (c) $\dfrac{1}{n}$

D3 힌트 위로 가는 제안은 $\pi(k+1)/\pi(k)=r<1$이라 확률 r로만 받아들여진다. 아래로 가는 제안은 언제나 받아들여진다.

답 (c) $1-\dfrac{r}{2}$ (d) $\dfrac{1-r}{2}$ (e) $\dfrac{2}{r}$

D4 힌트 (a) $\sum(x_i-\mu)^2=(n-1)s^2+n(\bar x-\mu)^2$을 쓰고 $u=1/\sigma^2$로 바꾸면 감마 적분이 된다. (b) 6회 정리 3.2의 t 밀도의 핵 $(1+t^2/\nu)^{-(\nu+1)/2}$과 맞춘다.

답 (c) 6.431 (d) 0.074