수리통계학 × 통계학의 역사

16회 워크북 — 풀이판

베이즈의 귀환
손으로 푸는 34문항

학생판과 같은 문제에 빈칸의 답과 완전한 풀이를 붙였다. 학생판으로 먼저 풀고, 힌트와 최종답으로도 막힐 때 이 판을 본다.

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.

풀이 정리 1.2(a)로 사후분포는 $\mathrm{Beta}(9,6)$이다. (a) $\frac{9}{15}=\frac{3}{5}$. (b) 밀도 $p^8(1-p)^5$의 로그를 미분해 0으로 두면 $\frac{8}{p}=\frac{5}{1-p}$, 곧 $p=\frac{8}{13}$이다. (c) $\frac{9\cdot 6}{15^2\cdot 16}=\frac{54}{3600}=\frac{3}{200}$. 사전평균은 $\frac{2}{5}$, 표본비율은 $\frac{7}{10}$이고 사후평균은 $\frac{5}{15}\cdot\frac{2}{5}+\frac{10}{15}\cdot\frac{7}{10}=\frac{2+7}{15}=\frac{3}{5}$로 둘의 가중평균이다(노트 연습 1.1).

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

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.

풀이 정리 1.2(b)로 사후분포는 $\Gamma(12,\ \text{율}\ 6)$이다. (a) $12/6=2$. (b) $P(X_6=0 \mid \mathbf{x})=E(e^{-\lambda}\mid \mathbf{x})=\int_0^\infty e^{-\lambda}\frac{6^{12}\lambda^{11}e^{-6\lambda}}{\Gamma(12)}d\lambda=\frac{6^{12}}{7^{12}}=(6/7)^{12}\approx 0.157$이다. 마지막 등호는 $\Gamma(12,\ \text{율}\ 7)$의 밀도가 적분하면 1이라는 것이다. MLE $\hat\lambda=2$를 넣으면 $e^{-2}\approx 0.135$로 더 작다. 사후예측은 λ의 불확실성까지 평균하므로, 볼록함수 $e^{-\lambda}$의 평균이 평균에서의 값보다 크다(2회 Jensen).

최종답 (a) 2 (b) $(6/7)^{12}\approx 0.157$

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?

풀이 정밀도는 $4+\frac{1}{4}=\frac{17}{4}$이다. (a) $\frac{4\cdot 2.5+\frac{1}{4}\cdot 0}{17/4}=\frac{40}{17}\approx 2.353$. (b) $\frac{4}{17}$. $\bar x$의 무게는 $\frac{4}{17/4}=\frac{16}{17}$이므로 $\bar x$를 0 쪽으로 $\frac{1}{17}$만큼 줄인다.

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

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 $L = (\theta - a)^2/\theta$.

풀이 (a) $\int_0^1 \theta \cdot 2\theta\, d\theta = \frac{2}{3}$. (b) $\int_0^m 2\theta\, d\theta = m^2 = \frac{1}{2}$에서 $m = 1/\sqrt{2} \approx 0.707$. (c) 밀도 2θ는 증가하므로 최빈값은 오른쪽 끝 1이다. (d) $E(w\theta \mid \mathbf{x}) = \int_0^1 \frac{1}{\theta}\,\theta \cdot 2\theta\, d\theta = 1$, $E(w \mid \mathbf{x}) = \int_0^1 \frac{1}{\theta} \cdot 2\theta\, d\theta = 2$이므로 $\frac{1}{2}$이다. 같은 사후분포에서 손실에 따라 0.5에서 1까지 다른 답이 나온다.

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

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.

풀이 (a) 예제 2.2의 공식에서 분자는 $16p(1-p)+(2-4p)^2 = 16p-16p^2+4-16p+16p^2 = 4$이고 분모는 $20^2 = 400$이므로 $R = \frac{4}{400} = \frac{1}{100}$이다. (b) $p(1-p)/16$은 $p = \frac{1}{2}$에서 최대 $\frac{1}{64}$이다. (c) $\frac{p(1-p)}{16} < \frac{1}{100} \iff p^2 - p + 0.16 > 0 \iff p < 0.2$ 또는 $p > 0.8$이므로 $p_0 = \frac{1}{5}$이다. 정리 2.3(b)로 δ는 minimax이고, 최대 위험이 $\frac{1}{64}$에서 $\frac{1}{100}$로 줄어드는 대신 끝 근처에서는 MLE에 진다.

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

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.

풀이 (a) 점수 $\frac{\partial}{\partial\theta}\log f = \frac{1}{\theta} - x$의 분산은 $\mathrm{Var}\,X = \frac{1}{\theta^2}$이므로 $I(\theta) = \theta^{-2}$, $\pi_J \propto \theta^{-1}$, $k = -1$이다. (b) 사후밀도 $\propto \theta^4 e^{-8\theta}$는 $\Gamma(5,\ \text{율}\ 8)$이므로 평균은 $\frac{5}{8}$이다. MLE 5/8과 같다. (c) $f = \frac{1}{\mu}e^{-x/\mu}$의 점수는 $-\frac{1}{\mu} + \frac{x}{\mu^2}$이고 분산은 $\frac{\mathrm{Var}\,X}{\mu^4} = \frac{1}{\mu^2}$이므로 $\pi_J \propto \mu^{-1}$, $k = -1$이다. 변수변환으로는 $\theta = 1/\mu$, $|d\theta/d\mu| = \mu^{-2}$이므로 $\theta^{-1}|d\theta/d\mu| = \mu \cdot \mu^{-2} = \mu^{-1}$로 같다.

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

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 $2P(S \ge 8)$과 비교하라.

풀이 $m_0 = \binom{10}{8}2^{-10} = \frac{45}{1024}$, $m_1 = \binom{10}{8}B(9,3) = 45 \cdot \frac{8!\,2!}{11!} = \frac{1}{11}$이다. (a) $B_{10} = \frac{1/11}{45/1024} = \frac{1024}{495} \approx 2.07$. (b) $10\log_{10} 2.07 \approx 3.16$ deciban. (c) $P(H_0 \mid \mathbf{x}) = \frac{1}{1+B_{10}} = \frac{495}{1519} \approx 0.326$. 유의확률은 $2 \cdot \frac{45+10+1}{1024} = 0.109$이다. 둘 다 H_0을 기각할 근거가 약하다고 말하지만, 사후확률 0.33은 유의확률 0.11보다 H_0에 훨씬 우호적이다.

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

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.

풀이 (a) $5+10+20 = 35$ deciban, 곧 가능도비 $10^{3.5} \approx 3162$. (b) 사후 오즈는 $-20+35 = 15$ deciban, 곧 $10^{1.5} \approx 31.6$이고 사후확률은 $\frac{31.6}{32.6} \approx 0.969$이다.

최종답 (a) 35 (b) 0.969

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.

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

풀이 $\pi_1 \cdot \frac{1}{2} = \pi_2 \cdot \frac{1}{4}$에서 $\pi_2 = 2\pi_1$, $\pi_2 \cdot \frac{1}{4} = \pi_3 \cdot \frac{1}{2}$에서 $\pi_3 = \pi_1$이다. 합이 1이므로 $\pi = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$이다. 상세 균형을 만족하는 π는 정리 4.3으로 정상분포다. 직접 확인하면 $\pi P = (\frac{1}{8}+\frac{1}{8},\ \frac{1}{8}+\frac{1}{4}+\frac{1}{8},\ \frac{1}{8}+\frac{1}{8}) = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$이다.

최종답 (a) $\frac{1}{4}$ (b) $\frac{1}{2}$

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?

풀이 (a) $\min\{1, \frac{9e^{-3}}{4e^{-2}}\} = \frac{9}{4e} \approx 0.828$. (b) $\min\{1, \frac{1\cdot e^{-1}}{4e^{-2}}\} = \frac{e}{4} \approx 0.680$. (c) $w(x)=\pi(x)/g(x) \propto x^2$이므로 $\min\{1, w(1)/w(2)\} = \frac{1}{4}$이다. 독립 제안은 목표의 x^2 인수 때문에 작은 값으로 가는 것을 더 자주 거절한다.

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

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.

풀이 조건부확률은 $p(Y=0 \mid X=0)=0.8$, $p(Y=0 \mid X=1)=0.4$, $p(X=1 \mid Y=0)=\frac{1}{3}$, $p(X=1 \mid Y=1)=\frac{3}{4}$이다. (a) $0.8\cdot\frac{1}{3}+0.2\cdot\frac{3}{4} = \frac{16}{60}+\frac{9}{60} = \frac{5}{12}$. (b) $0.4\cdot\frac{1}{3}+0.6\cdot\frac{3}{4} = \frac{8}{60}+\frac{27}{60} = \frac{7}{12}$. (c) $\pi_1 = \pi_0\cdot\frac{5}{12} + \pi_1\cdot\frac{7}{12}$에서 $\pi_0=\pi_1=\frac{1}{2}$이다. X의 주변분포도 $P(X=1)=0.2+0.3=\frac{1}{2}$로 같다(정리 4.7).

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

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).

풀이 보조정리로 $EZ^4 - \theta EZ^3 = 3EZ^2$이므로 $EZ^4 = \theta(\theta^3+3\theta) + 3(\theta^2+1) = \theta^4+6\theta^2+3$이다. $\theta=0$이면 3으로 표준정규의 넷째 적률이다.

최종답 $\theta^4+6\theta^2+3$

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?

풀이 $\|\mathbf{X}\|^2=1+4+0.25+2.25+1=8.5$이다. (a) $1-\frac{3}{8.5}=\frac{11}{17}$. (b) $\frac{11}{17}\cdot(-2)=-\frac{22}{17}\approx-1.29$. (c) $5-3^2\cdot\frac{1}{3}=2$로 $\mathbf{X}$의 위험 5보다 작다.

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

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.

풀이 (a) $\frac{8}{40}=\frac{1}{5}$. (b) $\frac{1}{1+\tau^2}=\frac{1}{5}$에서 $\tau^2=4$. (c) $B=\frac{1}{5}$이므로 $10\cdot\frac{4}{5}+2\cdot\frac{1}{5}=8.4$이다. Bayes 추정량의 위험 $p(1-B)=8$보다 0.4만 크고, $\mathbf{X}$의 위험 10보다 작다. τ^2을 몰라 치르는 값이 $2B=0.4$다.

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

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. — $\|u - v\|^2$의 전개

③ $\nabla \cdot \mathbf{g}(\mathbf{x}) = (p-2)^2/\|\mathbf{x}\|^2$이다. — $\partial_i(x_i/\|\mathbf{x}\|^2)$를 더한다

④ $E[(\mathbf{X} - \theta)'\mathbf{g}(\mathbf{X})] = (p-2)^2 E[1/\|\mathbf{X}\|^2]$이다. — 보조정리 5.1과 ③

⑤ $\|\mathbf{g}(\mathbf{X})\|^2 = (p-2)^2/\|\mathbf{X}\|^2$이다. — ①에서 바로

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. — $E\|\mathbf{X} - \theta\|^2 = p$

⑦ $p \geq 3$이면 이 기댓값은 양수이고 유한하므로 위험은 p보다 작다. — 연습 5.1, 5.3

풀이 빈칸의 빨간 글씨가 근거다. 노트 정리 5.2의 증명과 같다. ③의 합은 $\sum_i \frac{\|\mathbf{x}\|^2 - 2x_i^2}{\|\mathbf{x}\|^4} = \frac{p\|\mathbf{x}\|^2 - 2\|\mathbf{x}\|^2}{\|\mathbf{x}\|^4} = \frac{p-2}{\|\mathbf{x}\|^2}$이고, $\mathbf{g}$에는 상수 $p-2$가 하나 더 곱해져 있다. $p \leq 2$이면 $E[1/\|\mathbf{X}\|^2] = \infty$이고 ④의 보조정리 조건도 깨진다.

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. — 제안하고 받아들여야 옮긴다

② $\pi(x)k(x,y) = \min\{\pi(x)q(x,y),\ \pi(y)q(y,x)\}$이다. — 양수 c에 대해 $c\min\{a,b\} = \min\{ca, cb\}$

③ 따라서 $\pi(x)k(x,y) = \pi(y)k(y,x)$이다. — 오른쪽이 x, y에 대칭

④ $x = y$의 제자리 덧어리는 상세 균형을 어기지 않는다. — 양변이 같은 항

⑤ π는 정상분포다. — 정리 4.3

풀이 빈칸의 빨간 글씨가 근거다. 정규화 상수가 비에서 지워진다는 것은 α의 정의에서 이미 쓰였다. 상세 균형은 국소적인 조건이라 쌍마다 확인하면 되고, 정상분포는 정리 4.3이 준다.

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. — $\varphi'(u) = -u\varphi(u)$, 미적분의 기본정리

② $\int_0^\infty g'(z)\varphi(z)\,dz = \int_0^\infty u\varphi(u)\int_0^u g'(z)\,dz\,du$이다. — Fubini [T6], $E|g'(Z)| < \infty$

③ $= \int_0^\infty u\varphi(u)\big(g(u) - g(0)\big)du$이다. — g가 절대연속

④ 음의 쪽도 같아 $E[g'(Z)] = E[Z(g(Z) - g(0))]$이다. — $z < 0$에서 $\varphi(z) = -\int_{-\infty}^z u\varphi(u)du$

⑤ $= E[Zg(Z)]$이다. — $EZ = 0$

풀이 빈칸의 빨간 글씨가 근거다. ②의 Fubini를 정당화하는 절댓값의 이중적분은 $\int_0^\infty |g'(z)|\varphi(z)\,dz$이므로 가정이 바로 그 조건이다. 평균이 θ이면 $g(\cdot + \theta)$에 같은 증명을 쓴다.

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 X에서 함수 g를 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

풀이

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

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

③ $\nabla\cdot\mathbf{g} = (p-2)\sum_i \frac{\|\mathbf{x}\|^2 - 2x_i^2}{\|\mathbf{x}\|^4} = \frac{(p-2)^2}{\|\mathbf{x}\|^2}$. 몫의 미분

④ $E[(\mathbf{X}-\theta)'\mathbf{g}(\mathbf{X})] = E[\nabla\cdot\mathbf{g}(\mathbf{X})]$에는 θ가 없다. 보조정리 5.1

⑤ $R = p - 2(p-2)^2 E\|\mathbf{X}\|^{-2} + (p-2)^2 E\|\mathbf{X}\|^{-2} = p - (p-2)^2 E\|\mathbf{X}\|^{-2}$. $E\|\mathbf{X}-\theta\|^2 = p$

⑥ $0 < E\|\mathbf{X}\|^{-2} \leq \frac{1}{p-2} < \infty$이므로 $R < p$. 연습 5.1, 5.3

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

풀이 처음 다섯 단계는 B2의 풀이와 같다. 여섯째 단계: 연쇄가 $\pi(x) > 0$인 점에서 출발하면, $\pi(y) = 0$인 제안은 $\alpha(x,y) = 0$으로 거절되므로 그 뒤로도 $\pi > 0$인 점에만 있다. $\pi(x)q(x,y) = 0$이고 $\pi(y)q(y,x) > 0$인 쌍에서는 $\alpha(y,x) = \min\{1, 0\} = 0$이라 $\pi(x)k(x,y) = 0 = \pi(y)k(y,x)$이다. 양쪽이 모두 0이면 확인할 것이 없다.

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.

풀이 B4의 풀이와 같다. 조건: $\mathbf{g}$의 각 좌표는 원점을 지나지 않는 거의 모든 좌표 직선 위에서 매끄럽고, $|\partial_i g_i(\mathbf{x})| = (p-2)\frac{|\|\mathbf{x}\|^2 - 2x_i^2|}{\|\mathbf{x}\|^4} \leq \frac{3(p-2)}{\|\mathbf{x}\|^2}$이다. $\int_{\|\mathbf{x}\|\leq 1}\|\mathbf{x}\|^{-2}d\mathbf{x}$는 극좌표로 $c_p\int_0^1 r^{-2}r^{p-1}dr = c_p\int_0^1 r^{p-3}dr$이고 $p \geq 3$이면 유한하다. 정규밀도는 유계이므로 $E|\partial_i g_i(\mathbf{X})| < \infty$이다.

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.

풀이 $Z = \theta + W$, $W \sim \mathcal{N}(0,1)$, $h(w) = g(w+\theta)$로 두면 $h'(w) = g'(w+\theta)$이고 주장은 $E[Wh(W)] = E[h'(W)]$이다. 이것을 B3의 단계로 증명한다. Fubini의 조건: 양의 쪽에서 $\int_0^\infty \int_0^u |h'(z)|u\varphi(u)\,dz\,du = \int_0^\infty |h'(z)| \int_z^\infty u\varphi(u)\,du\,dz = \int_0^\infty |h'(z)|\varphi(z)\,dz \le E|h'(W)| < \infty$이다. 음의 쪽도 같다. 절대연속성은 $\int_0^u h' = h(u) - h(0)$에만 쓰인다.

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.

풀이 (a) $P(\phi \le t) = P(p \le \sqrt{t}) = \sqrt{t}$이므로 밀도는 $\frac{1}{2\sqrt{t}}$로, 0 근처에서 무한대로 간다. p에 "모름"을 균등으로 주면 p^2에는 작은 값을 선호하는 정보를 준다. (b) $\frac{\partial}{\partial \phi}\log f = \left(\frac{x}{\sqrt{\phi}} - \frac{1-x}{1-\sqrt{\phi}}\right)\frac{1}{2\sqrt{\phi}}$이다. 괄호의 분산은 $I_p(p) = \frac{1}{p(1-p)}$이므로 $I_\phi = \frac{1}{\sqrt{\phi}(1-\sqrt{\phi})} \cdot \frac{1}{4\phi}$이고 $\sqrt{I_\phi} \propto \phi^{-3/4}(1-\sqrt{\phi})^{-1/2}$이다. (c) $\text{Beta}(\frac{1}{2}, \frac{1}{2})$의 밀도 $\propto p^{-1/2}(1-p)^{-1/2}$를 $p = \sqrt{\phi}$로 옮기면 $\phi^{-1/4}(1-\sqrt{\phi})^{-1/2} \cdot \frac{1}{2\sqrt{\phi}} \propto \phi^{-3/4}(1-\sqrt{\phi})^{-1/2}$로 (b)와 같다. 정리 2.6의 한 예다.

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.

풀이 δ가 δ_0을 지배한다고 하자 [T2]. 그러면 $R(0,\delta) \le R(0,\delta_0) = 0$이므로 $E_0\delta(X)^2 = 0$, 곧 $P_0(\delta(X) \ne 0) = 0$이다. 모든 θ에서 $\mathcal{N}(\theta,1)$은 $\mathcal{N}(0,1)$과 같은 영집합을 갖는다(밀도가 어디서나 양수). 따라서 모든 θ에서 $\delta(X) = 0$ (확률 1)이고 $R(\theta,\delta) = R(\theta,\delta_0) = \theta^2$이다. 어떤 θ에서 엄격히 작다는 지배의 조건과 모순이다. 그러나 $\sup_\theta R(\theta,\delta_0) = \sup \theta^2 = \infty > 1 = \sup_\theta R(\theta, X)$이므로 minimax가 아니다. 허용성은 "모든 곳에서 지는 일은 없다"일 뿐이고, 한 점에 내기를 거는 추정량도 그 점에서 이기므로 허용적이다.

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?

풀이 (a) 임의의 δ'과 모든 k에서 $\sup_\theta R(\theta,\delta')\ge r(\pi_k,\delta')\ge r(\pi_k,\delta_k)$이다(평균은 최댓값을 넘지 않고, δ_k는 Bayes). $k\to\infty$이면 $\sup_\theta R(\theta,\delta')\ge c=\sup_\theta R(\theta,\delta)$이다 [T5]. (b) 정리 1.2(c)로 사후분산은 $\frac{1}{n+1/\tau^2}=\frac{\tau^2}{n\tau^2+1}$이고 자료에 의존하지 않으므로 Bayes 위험도 이 값이다(정리 2.1(a)). (c) $\tau\to\infty$이면 $r_\tau\to\frac{1}{n}$이고 $R(\mu,\bar X)=\frac{1}{n}$이므로 (a)로 $\bar X$는 minimax다. $\bar X$는 어떤 정칙 사전분포의 Bayes 추정량도 아니므로(명제 2.4) 정리 2.3(b)를 그대로 쓸 수 없다. 극한으로 그 빈자리를 메운다.

최종답 (b) $\frac{\tau^2}{n\tau^2+1}$

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?

풀이 (a) $a=\pi(x)q(x,y)$, $b=\pi(y)q(y,x)$로 두면 $r=b/a$이고 $\pi(x)k_B(x,y)=a\cdot\frac{b/a}{1+b/a}=\frac{ab}{a+b}$이다. 이것은 a와 b에 대칭이므로 $\pi(y)k_B(y,x)$와 같다. 제자리 덩어리는 정리 4.5의 증명과 같다. (b) $r\le1$이면 $\frac{r}{1+r}\le r$, $r>1$이면 $\frac{r}{1+r}<1$이다 [T4]. 받아들임 확률이 작으면 대각선 밖의 전이 확률이 작고, 같은 정상분포에서 더 오래 머문다(Peskun 1973). (c) $\frac{3}{1+3}=\frac{3}{4}$이다. MH는 1이다.

최종답 (c) $\frac{3}{4}$

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.

풀이 (a) $X = x$에서 Y를 뽑고 그 Y로 X'을 뽑으므로 전확률 공식(연속판)으로 $K(x,x') = \int \pi(y \mid x)\pi(x' \mid y)dy$이다. (b) $\pi_X(x)K(x,x') = \int \pi_X(x)\pi(y \mid x)\frac{\pi(x',y)}{\pi_Y(y)}dy = \int \frac{\pi(x,y)\pi(x',y)}{\pi_Y(y)}dy$이고, 이 식은 x와 x'을 바꿔도 같다. (c) 체계적 순서의 비가역성은 (X,Y) 쌍의 연쇄에 대한 것이다. 그 연쇄를 X에만 투영한 연쇄는 (b)대로 가역이다. 이 "자료 보강"(Tanner–Wong 1987) 구조에서는 Y가 X의 두 시점을 잇는 다리 역할만 한다.

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.)

풀이 (a) $E[Xg(X-1)] = \sum_{k\ge 1} g(k-1)e^{-\lambda}\frac{\lambda^k}{(k-1)!} = \lambda\sum_{j\ge 0} g(j)e^{-\lambda}\frac{\lambda^j}{j!} = \lambda E[g(X)]$이다. 절대수렴이 가정되어 첨자를 옮겨도 된다(2회 보조정리 0.1). (b) $E[\lambda(X+1)] = E[X\cdot X]$이므로 $EX^2 = \lambda(\lambda+1) = \lambda^2+\lambda$이고, 분산은 λ다.

최종답 (b) $\lambda^2+\lambda$

C7 임의의 점을 향한 축소 ★★

정리 5.2 · 20분

$\boldsymbol{\nu} \in \mathbb{R}^p$를 미리 정한 점이라 하자. $p \ge 3$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 $\boldsymbol{\theta}$는 어디인가?

풀이 $\mathbf{Y} = \mathbf{X}-\boldsymbol{\nu}$, $\boldsymbol{\eta} = \boldsymbol{\theta}-\boldsymbol{\nu}$로 두면 $\mathbf{Y} \sim \mathcal{N}_p(\boldsymbol{\eta}, I)$이고 $\delta_{\boldsymbol{\nu}} - \boldsymbol{\theta} = \big(1-\frac{p-2}{\|\mathbf{Y}\|^2}\big)\mathbf{Y} - \boldsymbol{\eta}$이다. 이것은 $\mathbf{Y}$에 대한 James–Stein의 오차이므로 정리 5.2로 위험은 $p-(p-2)^2E[1/\|\mathbf{Y}\|^2]$이다. $E[1/\|\mathbf{Y}\|^2] = E[1/(p-2+2K)]$, $K \sim \mathrm{Poi}(\|\boldsymbol{\eta}\|^2/2)$은 $\|\boldsymbol{\eta}\|$가 작을수록 크므로(노트 연습 5.3) 위험은 $\boldsymbol{\theta} = \boldsymbol{\nu}$에서 최소 2다. 줄일 곳을 참값 근처로 고르면 이득이 크고, 멀리 골라도 손해는 없다.

C8 축소 상수의 범위 ★★

정리 5.2 · 15분

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

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

풀이 (a) $\mathbf{g}=c\mathbf{X}/\|\mathbf{X}\|^2$이면 $\nabla\cdot\mathbf{g}=c(p-2)/\|\mathbf{X}\|^2$, $\|\mathbf{g}\|^2=c^2/\|\mathbf{X}\|^2$이므로 정리 5.2의 계산과 같이 $R=p-2c(p-2)E\|\mathbf{X}\|^{-2}+c^2E\|\mathbf{X}\|^{-2}$이다. (b) $E\|\mathbf{X}\|^{-2}\in(0,\infty)$이므로 $R<p$ (모든 θ) $\iff 2c(p-2)-c^2>0 \iff 0<c<2(p-2)$이다. $c=2(p-2)$이면 위험이 정확히 p로 지배하지 못한다. (c) 괄호 $2c(p-2)-c^2$은 $c=p-2$에서 최대이고 그 값은 $(p-2)^2$이다. $E\|\mathbf{X}\|^{-2}$이 c에 무관하므로 모든 θ에서 같은 c가 최선이다.

최종답 (c) $c=p-2$

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.)

풀이 (a) $P(\theta=1\mid x)=\frac{\varphi(x-1)}{\varphi(x-1)+\varphi(x+1)}=\frac{e^x}{e^x+e^{-x}}$이므로 $E(\theta\mid x)=\frac{e^x-e^{-x}}{e^x+e^{-x}}=\tanh x$이다. (b) $m(x)\propto e^{-x^2/2}(e^x+e^{-x})$이므로 $(\log m)'(x)=-x+\tanh x$이고 $x+(\log m)'(x)=\tanh x$로 같다. (c) $1-\tanh x=\frac{2e^{-2x}}{1+e^{-2x}}\approx 2e^{-2x}$로 지수적으로 1에 다가간다. 선형 축소(정규 사전분포)와 달리 큰 관측은 거의 줄이지 않는다.

최종답 $\tanh x$

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

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

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

풀이 (a) 정리 2.1(a)로 Bayes 추정량은 $E(e^{-\lambda}\mid \mathbf{x})$이고, 사후분포 $\Gamma(a+T,\ \text{율 } b+n)$에서

$$E(e^{-\lambda}\mid \mathbf{x}) = \int_0^\infty e^{-\lambda}\frac{(b+n)^{a+T}\lambda^{a+T-1}e^{-(b+n)\lambda}}{\Gamma(a+T)}d\lambda = \left(\frac{b+n}{b+n+1}\right)^{a+T}$$

이다. 마지막 등호는 율 $b+n+1$인 감마 밀도의 적분이 1이라는 것이다. (b) $(11/12)^{13} \approx 0.323$. (c) $0.9^{12} \approx 0.282$. MLE를 넣은 $e^{-1.2} \approx 0.301$은 둘 사이에 있다. (d) $\log \text{Bayes} = (a+T)\log(1-\frac{1}{b+n+1}) = -\frac{a+T}{b+n+1} + O_p(T/n^2) = -\frac{T}{n} + O_p(1/n)$이다. $T/n = O_p(1)$(7회 대수의 법칙)이므로 $\frac{a+T}{b+n+1} - \frac{T}{n} = \frac{an-bT-T}{n(b+n+1)} = O_p(1/n)$이다. UMVUE도 $T\log(1-\frac{1}{n}) = -\frac{T}{n} + O_p(1/n)$이다. 지수함수는 유계인 구간에서 립시츠이므로 세 값의 차이도 $O_p(1/n)$이다 [T8]. 사전분포와 불편성의 차이는 $1/n$ 크기이고 표준오차 $1/\sqrt{n}$보다 작다.

최종답 (a) $\left(\frac{b+n}{b+n+1}\right)^{a+T}$ (b) $(11/12)^{13} \approx 0.323$ (c) $0.9^{12} \approx 0.282$

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.

풀이 (a) 사후밀도 $\propto p^s(1-p)^{n-s}$에 w를 곱하면 $p^{s-1}(1-p)^{n-s-1}$, 곧 $\mathrm{Beta}(s,n-s)$의 핵이다. 정리 2.1(c)의 추정량 $E(wp\mid s)/E(w\mid s)$는 이 분포의 평균 $\frac{s}{s+n-s}=\frac{s}{n}$이다. (b) $s=0$이면 사후 기대손실은 $(n+1)\int_0^1 \frac{(p-a)^2}{p(1-p)}(1-p)^n\,dp$이다. $a>0$이면 $p\to 0$에서 피적분함수가 a^2/p처럼 커져 적분이 무한하고, $a=0$이면 $(n+1)\int_0^1 p(1-p)^{n-1}dp<\infty$이다. 따라서 Bayes 추정량은 $0=s/n$이다. $s=n$도 대칭이다. (c) $R(p,S/n)=\frac{E(S/n-p)^2}{p(1-p)}=\frac{p(1-p)/n}{p(1-p)}=\frac{1}{n}$. (d) Bayes 추정량의 위험이 상수이므로 정리 2.3(b)로 minimax다(Bayes 위험 $1/n$은 유한하다). 제곱손실에서는 $\delta_{\sqrt{n}/2,\sqrt{n}/2}$가 최대 위험을 더 줄였다. 손실을 끝에서 크게 매기면 끝을 잘 맞히는 MLE가 이긴다. minimax는 손실함수에 대한 주장이다.

최종답 (c) $\frac{1}{n}$

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.

풀이 (a) $\alpha(k,k+1)=\min\{1,r\}=r$, $\alpha(k+1,k)=\min\{1,1/r\}=1$이다. (b) $\pi(k)k(k,k+1)=(1-r)r^k\cdot\frac{1}{2}r$, $\pi(k+1)k(k+1,k)=(1-r)r^{k+1}\cdot\frac{1}{2}\cdot 1$로 같다. 두 칸 이상 떨어진 쌍은 양쪽이 0이다. (c) -1 제안(확률 $\frac{1}{2}$)은 거절되고, $+1$ 제안은 확률 $1-r$로 거절된다. $\frac{1}{2}+\frac{1}{2}(1-r)=1-\frac{r}{2}$. (d) 아래로는 언제나 옮기므로 $+1$ 제안의 거절만 남아 $\frac{1}{2}(1-r)$이다. (e) 0에서는 매 걸음 확률 $\frac{r}{2}$로 1로 옮겨 가고 아니면 0에 머문다. 걸음들은 독립이므로 처음 옮겨 가는 시점은

성공 확률 $\frac{r}{2}$인 기하분포이고 평균은 $\frac{2}{r}$이다(4회). $r=\frac{1}{2}$이면 4걸음이다.

최종답 (c) $1-\frac{r}{2}$ (d) $\frac{1-r}{2}$ (e) $\frac{2}{r}$

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^2은 표본평균과 표본분산이다.

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

(b) $\frac{\mu-\bar{x}}{s/\sqrt{n}}\,\big|\,\mathbf{x} \sim t_{n-1}$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.

풀이 (a) $A=(n-1)s^2+n(\mu-\bar{x})^2$로 두면 결합 사후밀도는 $(\sigma^2)^{-1}(\sigma^2)^{-n/2}e^{-A/(2\sigma^2)}$에 비례한다. $u=1/\sigma^2$, $d\sigma^2=u^{-2}du$로 바꾸면 $\int_0^\infty u^{n/2-1}e^{-Au/2}du=\Gamma(n/2)(2/A)^{n/2}\propto A^{-n/2}$이다 [T7]. (b) $t=\frac{\mu-\bar{x}}{s/\sqrt{n}}$이면 $A=(n-1)s^2\big(1+\frac{t^2}{n-1}\big)$이므로 밀도는 $\big(1+\frac{t^2}{n-1}\big)^{-n/2}$에 비례하고, 이것은 자유도 $\nu=n-1$의 t 밀도의 핵이다(6회 정리 3.2). $\mu\mapsto t$는 일차변환이라 야코비안은 상수다. (c) $5+2.2622\cdot\frac{2}{\sqrt{10}}\approx 6.431$이다. 6회 정리 5.2의 t-구간과 숫자까지 같다. (d) $t=\frac{6-5}{2/\sqrt{10}}=1.581$이므로 $P(t_9>1.581)\approx 0.074$이다. 빈도주의의 한쪽 유의확률과 같은 숫자지만, 여기서는 μ에 대한 확률로 읽는다.

최종답 (c) 6.431 (d) 0.074