수리통계학 × 통계학의 역사

12회 워크북

최적 검정과 결정 손으로 푸는 34문항

강의노트 12회를 읽은 뒤 손으로 푼다. 이번 회차의 최적성은 모두 한 줄의 부등식 $(\varphi^* - \varphi)(f_1 - kf_0) \geq 0$에서 나온다. 그 부등식을 빈 종이에서 다시 세우고, 임계값을 손으로 계산하는 것이 이 워크북의 일이다. 문제 아래의 줄 친 공간에 직접 쓴다.

부	하는 일	문항	예상 시간
A 계산 드릴	최강력·UMP 검정의 임계값, 가능도비, SPRT 경계, Bayes 임계값	14	2시간 6분
B 빈칸 증명	노트의 단계형 증명을 책을 덮고 다시 쓴다	7	2시간 30분
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 `12_optimal_tests_notes.pdf` — 문항의 "정리 2.4" 같은 번호는 노트의 번호다

도구 `00_toolkit_notes.pdf` — [T1] 지시함수, [T2] 대우·귀류, [T4] 부등식, [T6] 합과 적분의 순서 바꾸기

검산 노트북 `12_optimal_tests.ipynb`의 "워크북 검산" 절: `check("A2b", ` 손으로 구한 답`)`

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. 정규분포표, χ^2 표에서 읽은 값은 유효숫자 셋째 자리까지 맞으면 맞다고 본다.
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. 표준정규분포의 분포함수를 Φ, 상위 α 분위수를 z_α로 쓴다($z_{0.05} = 1.645$, $z_{0.10} = 1.282$).

A

계산 드릴

최강력 검정과 균일최강력 검정의 임계값, 가능도비 통계량, SPRT의 경계, Bayes 임계값을 손으로 계산한다. 답이 수나 식으로 떨어지므로 다 풀면 뒤의 최종답이나 노트북의 `check()`로 맞춰 본다. 정규분포표와 χ^2 표를 써도 된다.

A1 베르누이의 최강력 검정 ★

정리 1.2 · 10분

$X_1, \dots, X_{10} \overset{\text{iid}}{\sim} \text{Bernoulli}(p)$로 $H_0 : p = 0.2$ 대 $H_1 : p = 0.5$를 크기 0.05로 검정한다. $T = \sum X_i$이다.

(a) 확률화하지 않은 검정 "$T \geq c$이면 기각" 가운데 크기가 0.05 이하인 것으로 가장 작은 c를 구하라.

(b) $T = c - 1$에서 확률 γ로 기각해 크기를 정확히 0.05로 맞추는 γ를 구하라.

(c) (b)의 확률화 검정의 $p = 0.5$에서의 검정력을 구하라.

A2 지수분포 관측 하나 ★

정리 1.2 · 8분

X의 밀도가 $\theta e^{-\theta x}$ $(x > 0)$이다. 관측 하나로 $H_0 : \theta = 1$ 대 $H_1 : \theta = 3$을 크기 0.1로 검정한다. (a) 최강력 검정의 기각역이 $\{X < c\}$임을 보이고 c를 구하라. (b) 검정력을 구하라.

A3 Poisson 관측 하나 ★★

정리 1.2 · 10분

$X \sim \mathrm{Poi}(\lambda)$ 하나로 $H_0 : \lambda = 2$ 대 $H_1 : \lambda = 5$를 크기 0.05로 검정한다.

(a) 확률화하지 않은 검정 "$X \geq c$이면 기각" 가운데 크기가 0.05 이하인 가장 작은 c와 (b) 그 크기를 구하라.

(c) 크기를 정확히 0.05로 맞추는 확률화 상수 γ ($X = c-1$에서)를 구하라.

(d) 확률화 검정의 $\lambda = 5$에서의 검정력을 구하라.

A4 검정력을 위한 표본 크기 ★

정리 2.4 · 8분

$X_i \overset{\text{iid}}{\sim} \mathcal{N}(\mu, 4)$로 $H_0 : \mu \leq 0$ 대 $H_1 : \mu > 0$을 크기 0.05의 UMP 검정으로 검정한다. $\mu = 1$에서 검정력이 0.9 이상이 되는 가장 작은 n을 구하라.

A5 베타형 밀도의 UMP ★

정의 2.2, 정리 2.4 · 10분

X의 밀도가 $f(x;\theta) = \theta x^{\theta-1}$ $(0 < x < 1,\ \theta > 0)$이다. 관측 하나로 검정한다.

(a) 이 족이 x에 대해 단조가능도비를 가짐을 보여라.

(b) $H_0 : \theta \leq 1$ 대 $H_1 : \theta > 1$의 크기 0.05 UMP 검정 "$X > c$이면 기각"의 c를 구하라.

(c) $\theta = 2$에서의 검정력을 구하라.

A6 지수 평균의 한쪽 검정 ★★

정리 2.4 · 10분

$X_1, \dots, X_{10} \overset{\text{iid}}{\sim}$ 지수분포(평균 θ)로 $H_0 : \theta \le 2$ 대 $H_1 : \theta > 2$를 크기 0.05로 검정한다. (a) UMP 검정 "$\sum X_i > c$이면 기각"의 c를 χ^2 분포의 분위수로 구하라. (b) $\theta = 4$에서의 검정력을 구하라.

A7 정규 평균의 가능도비 ★

정의 3.1 · 6분

$X_i \overset{\text{iid}}{\sim} \mathcal{N}(\mu, 1)$, $n = 16$, $\bar{x} = 0.6$이다. $H_0 : \mu = 0$ 대 $H_1 : \mu \neq 0$의 (a) $-2\log\lambda$를 구하고, (b) $\chi^2(1)$로 읽은 p-값을 구하라. 양쪽 z 검정의 p-값과 비교하라.

A8 이항 비율의 가능도비 ★

정의 3.1 · 8분

$X \sim \mathrm{b}(100, p)$에서 $x = 60$을 관측했다. $H_0 : p = 0.5$의 가능도비 통계량 $-2\log\lambda$를 구하라.

A9 Wald의 경계 ★

정리 4.2 · 6분

SPRT의 목표 오류율을 $\alpha = 0.01$, $\beta = 0.05$로 둔다. (a) 위 경계 A, (b) 아래 경계 B를 구하고, (c) 실제 1종 오류율 α'의 상한을 정리 4.2로 구하라.

A10 정규 SPRT의 평균 표본수 ★★

정리 4.2, 보조정리 4.3 · 12분

$X_i \overset{\text{iid}}{\sim} \mathcal{N}(\mu, 1)$로 $H_0 : \mu = 0$ 대 $H_1 : \mu = 0.5$를 $\alpha = \beta = 0.05$의 SPRT로 검정한다. $Z = \log \frac{f_1(X)}{f_0(X)}$이다. (a) $E_0 Z$를 구하라. (b) 넘침을 무시한 Wald의 근사로 $E_0 N$을 구하라. (c) 같은 α, β를 내는 고정 표본의 한쪽 z 검정에 필요한 가장 작은 n을 구하라.

A11 비용이 다른 Bayes 검정 ★

따름정리 5.3 · 8분

관측 하나 x로 $H_0 : \mathcal{N}(0,1)$ 대 $H_1 : \mathcal{N}(1,1)$을 결정한다. 사전확률은 $\pi_0 = 0.9$, $\pi_1 = 0.1$이고, 1종 오류의 비용은 1, 2종 오류의 비용은 4이다. (a) 가능도비 임계값 k와 (b) x의 기각 경계를 구하라.

A12 두 무리의 선형판별 ★★

따름정리 5.5 · 12분

$\boldsymbol{\mu}_0 = (0,0)$, $\boldsymbol{\mu}_1 = (2,1)$, $\Sigma = \left(\begin{smallmatrix} 1 & 0.5 \\ 0.5 & 1 \end{smallmatrix}\right)$, $\pi_0 = \pi_1$이다. (a)(b) 판별 방향 $\mathbf{w} = \Sigma^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)$의 두 성분, (c) Mahalanobis 거리 Δ, (d) 오분류 확률을 구하라.

A13 정규 점수의 ROC ★

8.1절 · 8분

음성의 점수가 $\mathcal{N}(0,1)$, 양성의 점수가 $\mathcal{N}(1,1)$이다. (a) 거짓양성률 0.05에서의 참양성률과 (b) ROC 곡선 아래 넓이를 구하라.

A14 Lindley의 역설, 숫자로 ★★

예제 7.2 · 10분

$X_i \overset{\text{iid}}{\sim} \mathcal{N}(\mu, 1)$, $n = 400$에서 $z = \sqrt{n}\bar{x} = 2.5$가 나왔다. H_1 아래 $\mu \sim \mathcal{N}(0,1)$로 둔다. (a) 양쪽 p-값, (b) Bayes 인수 B_{01}, (c) 두 가설의 사전확률이 같을 때 $P(H_0 \mid \mathbf{x})$를 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 결과, 도구 번호를 쓴다.

B1 Neyman–Pearson의 (b)(c), 근거 채우기 ★★

① · 정리 1.2 · 15분

φ^*는 $f_1 > kf_0$에서 1, $f_1 < kf_0$에서 0인 검정이고 $k \geq 0$이다. 각 단계의 근거를 써라.

① 크기가 $E_{\theta_0}\varphi^*$ 이하인 검정 φ에 대해, 모든 $\mathbf{x}$에서 $(\varphi^* - \varphi)(f_1 - kf_0) \geq 0$이다.

왜? _____

② 적분한다.

$$0 \leq \int (\varphi^* - \varphi)(f_1 - kf_0)\, d\mathbf{x}$$ 왜? _____

$$= \left(E_{\theta_1}\varphi^* - E_{\theta_1}\varphi\right) - k\left(E_{\theta_0}\varphi^* - E_{\theta_0}\varphi\right)$$ 왜? _____

③ $E_{\theta_1}\varphi^* - E_{\theta_1}\varphi \geq k\left(E_{\theta_0}\varphi^* - E_{\theta_0}\varphi\right) \geq 0$이다. 왜? _____

④ φ도 크기 α의 최강력 검정이면 ②의 적분은 정확히 0이다.

왜? _____

⑤ $f_1 \neq kf_0$인 곳에서는 거의 어디서나 $\varphi = \varphi^*$이다. 왜? _____

B2 Karlin–Rubin, 근거 채우기 ★★

① · 정리 2.4 · 15분

분포족이 T에 대해 단조가능도비를 갖고, φ^*는 $T>c$에서 1, $T<c$에서 0, $E_{\theta_0}\varphi^*=\alpha$이다. 각 단계의 근거를 써라.

① $\theta'<\theta''$이면 $f(\mathbf{x};\theta'')/f(\mathbf{x};\theta')=g(T(\mathbf{x}))$, g는 감소하지 않는다. $k=g(c)$로 둔다.

왜? _____

② $f''>kf'$이면 $T>c$이므로 $\varphi^*=1$이고, $f''<kf'$이면 $T<c$이므로 $\varphi^*=0$이다.

왜? _____

③ φ^*는 θ' 대 θ''의 자기 크기 최강력 검정이다. 왜? _____

④ $E_{\theta''}\varphi^*\geq E_{\theta'}\varphi^*$이다. 왜? _____

⑤ $\sup_{\theta\leq\theta_0}E_\theta\varphi^*=E_{\theta_0}\varphi^*=\alpha$이다. 왜? _____

⑥ $\theta_1>\theta_0$이고 φ의 크기가 α 이하이면 $E_{\theta_0}\varphi\leq\alpha$이다. 왜? _____

⑦ $E_{\theta_1}\varphi^*\geq E_{\theta_1}\varphi$이고, θ_1이 아무것이므로 UMP다. 왜? _____

B3 Wald의 등식, 근거 채우기 ★★

① · 보조정리 4.3 · 15분

$Z_1,Z_2,\ldots$가 iid, $E|Z_1|<\infty$, N이 멈춤시각, $EN<\infty$이다. 각 단계의 근거를 써라.

① $\sum_{i=1}^{N}Z_i=\sum_{i=1}^{\infty}Z_i\,1\{N\geq i\}$이다. 왜? _____

② $\{N\geq i\}$는 $Z_1,\ldots,Z_{i-1}$로 정해지므로 Z_i와 독립이다. 왜? _____

③ $E\big[Z_i 1\{N\geq i\}\big]=EZ_1\,P(N\geq i)$이다. 왜? _____

④ $E\sum_{i\leq N}Z_i=\sum_i EZ_1\,P(N\geq i)=EZ_1\cdot EN$이다. 왜? _____

조건 확인: ④의 첫 등호에서 합과 기댓값의 순서를 바꿀 수 있는 이유를 쓰라.

_____

B4 Neyman–Pearson, 단계의 이름만 ★★

② · 정리 1.2 · 25분

정리 1.2의 (a)(b)(c)를 증명하라. 단계의 이름만 주어진다. 각 단계 아래에 식과 근거를 써라.

① **(a) 가능도비의 생존함수 $F(k)$를 정의하고 성질을 적는다**

② 임계값 k를 고른다

③ 경계에서 확률화해 크기를 맞춘다

④ (b) 모든 점에서 성립하는 부등식

⑤ 적분한다

⑥ 결론

⑦ (c) 적분이 0이 되는 이유

⑧ 거의 어디서나의 결론

B5 Karlin–Rubin, 단계의 이름만 ★★

② · 정리 2.4 · 25분

정리 2.4를 증명하라. 단계의 이름만 주어진다.

① 두 모수값을 고정하고 단조가능도비를 쓴다

② 기각역이 가능도비의 위쪽 집합임을 보인다

③ 그래서 Neyman–Pearson 검정이다

④ 검정력 함수의 단조성

⑤ 크기

⑥ 비교할 검정을 고른다

⑦ 균일최강력의 결론

B6 Neyman–Pearson, 전략만 ★★

③ · 정리 1.2 · 30분

정리 1.2의 (a)(b)(c)를 증명하라.

전략 (a)는 가능도비의 생존함수에서 임계값을 고르고 경계에서 확률화한다. (b)는 모든 점에서 성립하는 부등식 하나를 적분한다. (c)는 그 부등식이 등호가 되는 조건을 읽는다.

B7 Wald의 등식, 전략만 ★★

③ · 보조정리 4.3 · 25분

보조정리 4.3을 증명하라. 그리고 N이 멈춤시각이 아니면 몇 번째 단계가 무너지는지 밝혀라.

전략 무작위 개수의 합을 지시함수로 고정된 무한합으로 바꾼다. 멈춤시각이라는 조건이 각 항의 두 인수를 독립으로 만든다.

C

증명 문제

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

C1 하나의 검정이 여러 대립가설에 최강력 ★★

[T2] · 정리 1.2 · 15분

φ^*가 f_0 대 f_1의 크기 α Neyman–Pearson 검정(임계값 $k>0$)이다. 다른 밀도 f_2가 있어 $f_0>0$인 곳에서 $f_2(\mathbf{x})/f_0(\mathbf{x}) = \psi\big(f_1(\mathbf{x})/f_0(\mathbf{x})\big)$이고 ψ가 순증가함수이다.

(a) $\{f_1 > k f_0\} = \{f_2 > \psi(k) f_0\}$, $\{f_1 < k f_0\} = \{f_2 < \psi(k) f_0\}$임을 보여라.

(b) φ^*가 f_0 대 f_2의 크기 α 최강력 검정임을 보여라.

(c) 이것이 정리 2.4의 증명의 어느 단계와 같은 생각인지 밝혀라.

C2 단조가능도비와 기댓값의 단조성 ★★

[T2] · 정의 2.2 · 20분

분포족이 T에 대해 단조가능도비를 갖고, ψ가 감소하지 않는 함수이며 $E_\theta|\psi(T)|<\infty$이다. $\theta_1<\theta_2$이면 $E_{\theta_1}\psi(T)\le E_{\theta_2}\psi(T)$임을 보여라. $f_j=f(\cdot;\theta_j)$로 쓴다.

(a) $A=\{f_2>f_1\}$, $B=\{f_2<f_1\}$이면 $\mathbf{x}\in A$, $\mathbf{y}\in B$에 대해 $T(\mathbf{x})>T(\mathbf{y})$임을 보여라.

(b) $a=\inf_A\psi(T)$, $b=\sup_B\psi(T)$로 두면 $a\ge b$이고, $D=\int_A(f_2-f_1)=\int_B(f_1-f_2)$임을 보여라.

(c) $E_{\theta_2}\psi(T)-E_{\theta_1}\psi(T)\ge(a-b)D\ge 0$을 보여라.

C3 균등분포에는 양쪽 UMP가 있다 ★★★

[T2] · 정리 1.2 · 25분

$X_1,\dots,X_n\overset{\text{iid}}{\sim}\mathrm{U}(0,\theta)$로 $H_0:\theta=1$ 대 $H_1:\theta\neq 1$을 검정한다. 검정 φ는 $X_{(n)}>1$이거나 $X_{(n)}\le\alpha^{1/n}$이면 기각한다.

(a) φ의 크기가 α임을 보여라.

(b) $\theta>1$이면 검정력이 $1-(1-\alpha)\theta^{-n}$이고, 이것이 1 대 θ의 크기 α 최강력 검정의 검정력과 같음을 보여라.

(c) $\theta<1$이면 크기 α인 어떤 검정의 검정력도 $\min(1,\alpha\theta^{-n})$을 넘지 못함을 보이고, φ가 그 값을 얻음을 보여라.

(d) φ가 UMP임을 결론짓고, $n=5$, $\alpha=0.05$, $\theta=0.8$에서의 검정력을 구하라.

C4 Wilks 정리가 맞지 않는 가능도비 ★★

[T2] · 정의 3.1, 9회 정리 5.2 · 20분

$X_1, \dots, X_n \overset{\text{iid}}{\sim} \mathrm{U}(0, \theta)$로 $H_0 : \theta = \theta_0$을 가능도비로 검정한다.

(a) $\lambda(\mathbf{x}) = (x_{(n)}/\theta_0)^n$ $(x_{(n)} \le \theta_0)$, 0 $(x_{(n)} > \theta_0)$임을 보여라.

(b) H_0 아래 $-2\log\lambda \sim \chi^2(2)$임을 보여라. 자유도는 n에 무관하다.

(c) Wilks 정리를 잘못 적용해 $\chi^2(1)$의 임계값 3.841로 기각하면 실제 크기는 얼마인가?

(d) Wilks 정리의 어느 가정이 깨졌는가?

C5 Wald의 둘째 등식 ★★

[T1] · 보조정리 4.3 · 20분

$Z_1, Z_2, \dots$ 가 iid이고 $EZ_1 = 0$, $EZ_1^2 = \sigma^2$ 이다. $S_n = \sum_{i \le n} Z_i$ 이다.

(a) N이 유계인 멈춤시각($N \le m$)이면 $ES_N^2 = \sigma^2 EN$임을 보여라.

(b) 이 등식이 $EN < \infty$인 멈춤시각에서도 성립함을 인정하자. 대칭 ± 1 걸음 S_n이 $(-a, b)$를 처음 벗어나는 시각 N에 대해, Wald의 등식으로 $P(S_N = b)$를 구하고 (a)로 $EN = ab$를 보여라. $a = 3$, $b = 5$이면 EN은?

C6 Bayes 규칙은 허용적이다 ★★

[T2] · 정의 5.1, 정리 5.2 · 15분

모수공간이 유한하다: $\Omega = \{\theta_1, \dots, \theta_m\}$. 사전분포가 모든 점에 양의 확률 $\pi_j > 0$을 주고, δ_π가 Bayes 규칙이며 $r(\pi, \delta_\pi) < \infty$이다. 어떤 규칙 δ'도 δ_π를 "지배" 할 수 없음을 보여라. 곧 모든 j에서 $R(\theta_j, \delta') \le R(\theta_j, \delta_\pi)$이고 어떤 j에서 부등호가 엄격한 δ'는 없다. 사전확률이 0인 점이 있으면 어디서 증명이 막히는가?

C7 위험이 일정한 Bayes 규칙은 minimax다 ★★

정의 5.1 · 15분

δ가 사전분포 π의 Bayes 규칙이고 모든 θ에서 $R(\theta,\delta)=c$ (상수)이다.

(a) 모든 규칙 δ'에 대해 $\sup_\theta R(\theta,\delta') \ge r(\pi,\delta') \ge r(\pi,\delta) = c = \sup_\theta R(\theta,\delta)$임을 보이고, δ가 minimax임을 결론지어라.

(b) 노트 연습 5.4의 검정 "$x > \frac{1}{2}$이면 기각"이 (a)의 예임을 확인하라. 위험은 두 오류 확률이다.

C8 확률화하지 않으면 가능도비 순서가 깨진다 ★★

정리 1.2 · 20분

결과가 네 개인 실험에서 두 가설의 확률이 다음과 같다. 크기 0.05로 검정한다.

결과	x_1	x_2	x_3	x_4
f_0	0.04	0.03	0.02	0.91
f_1	0.08	0.054	0.035	0.831

(a) 각 결과의 가능도비를 구하고 큰 순서로 늘어놓아라.

(b) 확률화하지 않은 검정(결과의 부분집합) 가운데 크기 0.05 이하에서 검정력이 가장 큰 것을 모든 부분집합을 따져 찾고, 그 검정력을 구하라. 가능도비가 가장 큰 결과를 포함하는가?

(c) 정리 1.2의 확률화 검정의 검정력을 구하라.

C9 (크기, 검정력)의 집합 ★★

[T2] · 정리 1.2 · 20분

f_0 대 f_1의 모든 검정 φ에 대해 점 $\big(E_0\varphi,\, E_1\varphi\big)$의 집합을 S라 하자.

(a) S가 볼록집합이고 $(0,0)$, $(1,1)$을 포함함을 보여라.

(b) $(\alpha,\gamma)\in S$이면 $(1-\alpha,1-\gamma)\in S$임을 보여라.

(c) S의 위쪽 경계 $\gamma^*(\alpha)=\max\{\gamma:(\alpha,\gamma)\in S\}$가 Neyman–Pearson 검정의 검정력이고, α의 오목함수이며 $\gamma^*(\alpha)\ge\alpha$임을 보여라.

D

퀄형

여러 절과 앞 회차의 도구를 잇는 종합 문제다. 머리의 제한 시간을 재고, 노트를 덮고 푼다. 검정을 유도할 때는 가능도비가 어떤 통계량의 단조함수인지부터 밝힌다.

D1 이동한 지수분포의 UMP 검정 ★★★

정리 2.4, 8회 정리 1.2 · 제한 35분

$X_1, \dots, X_n$이 iid이고 밀도가 $f(x;\theta) = e^{-(x-\theta)}$ $(x > \theta)$이다.

(a) 결합밀도를 써서 이 족이 $X_{(1)}$에 대해 단조가능도비를 가짐을 보여라(가능도비는 ∞를 허용한다).

(b) $X_{(1)} - \theta$가 비율 n의 지수분포임을 보이고, $H_0 : \theta \le 0$ 대 $H_1 : \theta > 0$의 크기 α UMP 검정이 "$X_{(1)} > c$이면 기각", $c = -\frac{1}{n}\log\alpha$임을 보여라.

(c) 검정력 함수를 구하라. $\theta > c$이면 검정력은 얼마인가?

(d) $n = 10$, $\alpha = 0.05$이면 c와 $\theta = 0.2$에서의 검정력은?

D2 두 Poisson 비율의 비교 ★★★ 정의 3.1, 11회 정리 1.3 · 제한 35분

같은 기간 동안 두 공장의 사고 수 $X \sim \mathrm{Poi}(\lambda_1)$, $Y \sim \mathrm{Poi}(\lambda_2)$를 관측한다($X$, Y 독립). $H_0 : \lambda_1 = \lambda_2$를 검정한다.

(a) 가능도비 통계량이 $-2\log\lambda = 2\big[x\log\frac{2x}{x+y} + y\log\frac{2y}{x+y}\big]$임을 보이고 $x = 12$, $y = 4$에서 계산하라.

(b) H_0 아래 $X + Y = n$이 주어지면 $X \sim \mathrm{b}(n, \frac{1}{2})$임을 보이고, 이 조건부분포가 공통 λ에 의존하지 않는 이유를 충분성으로 설명하라.

(c) $x = 12$, $y = 4$에서 조건부 정확 검정의 한쪽 p-값 $P(X \geq 12 \mid X + Y = 16)$을 구하라.

D3 불량률의 축차 검사 ★★★

정리 4.2, 보조정리 4.3 · 제한 35분

불량률 p인 제품을 하나씩 검사해 $H_0 : p = 0.1$ 대 $H_1 : p = 0.3$을 $\alpha = \beta = 0.05$의 SPRT로 결정한다. n개까지의 불량품 수를 d_n이라 하자.

(a) 계속할 영역이 $sn - h < d_n < sn + h$ 꼴의 띠임을 보이고 (a) 기울기 s와 (b) 절편 h를 구하라.

(c) Wald의 근사로 (c) $E_{0.1}N$과 (d) $E_{0.3}N$을 구하라.

(d) 양품만 계속 나오면 (e) 몇 개째에 H_0를 택하는가? 불량품만 계속 나오면 (f) 몇 개째에 H_1을 택하는가?

D4 비용이 다른 분류: Bayes와 minimax ★★★

정리 5.2, 따름정리 5.3 · 제한 35분

관측 x 하나로 무리 0 ($\mathcal{N}(0,1)$)과 무리 1 ($\mathcal{N}(2,1)$)을 가른다. 무리 1의 사전확률은 0.2이고, 무리 1을 무리 0으로 잘못 부르는 비용은 5, 반대 오류의 비용은 1이다(맞으면 0). 규칙은 "$x > t$이면 무리 1"이다.

(a) Bayes 규칙의 t를 구하라.

(b) 그 Bayes 위험 $0.8 \cdot 1 \cdot P_0(X > t) + 0.2 \cdot 5 \cdot P_1(X \le t)$를 구하라.

(c) 두 무리의 위험 $1 \cdot P_0(X > t)$와 $5 \cdot P_1(X \le t)$를 같게 하는 t(minimax 규칙)를 수치로 구하라.

(d) (c)의 규칙이 minimax인 이유를 C7로 설명하라. 어떤 사전확률에 대한 Bayes 규칙인가?

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 가능도비 $\left(\frac{0.5}{0.2}\right)^{T}\left(\frac{0.5}{0.8}\right)^{10-T}$는 T의 증가함수다. H_0 아래 $P(T \ge 5) = 0.0328$, $P(T = 4) = 0.0881$이다.

답 (a) $c = 5$ (b) $\gamma \approx 0.195$ (c) ≈ 0.663

A2 힌트 가능도비 $3e^{-2x}$는 x의 감소함수다. 그래서 "가능도비가 크다"는 "x가 작다"이다.

답 (a) $c = \log\frac{10}{9} \approx 0.105$ (b) 0.271

A3 힌트 $\lambda = 2$에서 $P(X \ge 5) = 0.0527$이다. 0.05보다 조금 크다. $P_2(X = 5) = e^{-2}\cdot 2^5/5!$이다.

답 (a) $c = 6$ (b) ≈ 0.0166 (c) $\gamma \approx 0.926$ (d) ≈ 0.547

A4 힌트 UMP 검정은 $\bar X > 1.645\cdot 2/\sqrt{n}$이다. $\mu = 1$에서 $P(\bar X > c) = \Phi\big(\sqrt{n}(1-c)/2\big)$이 0.9 이상이려면 $\sqrt{n}/2 \ge 1.645 + 1.282$.

답 $n = 35$

A5 힌트 $\theta_1 < \theta_2$이면 가능도비는 $\frac{\theta_2}{\theta_1}x^{\theta_2-\theta_1}$이다. $\theta = 1$이면 $X \sim \mathrm{U}(0,1)$이다.

답 (b) $c = 0.95$ (c) 0.0975

A6 힌트 평균 θ인 지수분포 n개의 합은 감마(n,θ)이고, $\frac{2}{\theta}\sum X_i \sim \chi^2(2n)$이다(4회). $\theta = 2$이면 $\sum X_i$ 자체가 $\chi^2(20)$이다.

답 (a) $c = \chi^2_{20,\,0.05} \approx 31.41$ (b) ≈ 0.735

A7 힌트 $\sup_\mu L$은 $\mu = \bar x$에서 얻는다. $-2\log\lambda = n\bar x^2 = z^2$이다.

답 (a) 5.76 (b) ≈ 0.0164

A8 힌트 $\hat p = 0.6$이다. $-2\log\lambda = 2\big[x\log\frac{\hat p}{p_0} + (n-x)\log\frac{1-\hat p}{1-p_0}\big]$.

답 ≈ 4.03

A9 힌트 $A = \log\frac{1-\beta}{\alpha}$, $B = \log\frac{\beta}{1-\alpha}$, $\alpha' \le \frac{\alpha}{1-\beta}$.

답 (a) $A = \log 95 \approx 4.554$ (b) $B = \log\frac{5}{99} \approx -2.986$ (c) $\alpha' \le \frac{1}{95} \approx 0.0105$

A10 힌트 $Z = \delta X - \delta^2/2$이다. $E_0 N \approx \frac{(1-\alpha)B + \alpha A}{E_0 Z}$. 고정 표본은 $\sqrt{n}\,\delta \geq z_{0.05} + z_{0.05}$.

답 (a) $E_0 Z = -\frac{1}{8}$ (b) ≈ 21.2 (c) $n = 44$

A11 힌트 $k = \frac{\pi_0 c_{\mathrm{I}}}{\pi_1 c_{\mathrm{II}}}$이고 $f_1(x)/f_0(x) = e^{x - 1/2}$이다.

답 (a) $k = \frac{9}{4}$ (b) $x > \frac{1}{2} + \log\frac{9}{4} \approx 1.311$

A12 힌트 $\Sigma^{-1} = \frac{1}{0.75}\begin{pmatrix} 1 & -0.5 \\ -0.5 & 1 \end{pmatrix}$. $\Delta^2 = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)'\mathbf{w}$.

답 (a) $w_1 = 2$ (b) $w_2 = 0$ (c) $\Delta = 2$ (d) $\Phi(-1) \approx 0.159$

A13 힌트 (b) AUC $= P(S_1 > S_0)$이고 $S_1 - S_0 \sim \mathcal{N}(1, 2)$이다.

답 (a) ≈ 0.260 (b) $\Phi(1/\sqrt{2}) \approx 0.760$

A14 힌트 예제 7.2의 식 $B_{01} = \sqrt{1+n}\exp\!\left(-\frac{z^2}{2}\cdot\frac{n}{n+1}\right)$을 쓴다. $P(H_0 \mid \mathbf{x}) = \frac{B_{01}}{1+B_{01}}$.

답 (a) ≈ 0.0124 (b) $B_{01} \approx 0.887$ (c) ≈ 0.470

B1 힌트 근거는 짧게 쓴다. ①은 $f_1 > kf_0$, $f_1 < kf_0$, $f_1 = kf_0$의 세 경우로 나눠 확인한 것이다.

B2 힌트 ④와 ⑦은 모두 ③의 결과를 쓴다. 하나는 따름정리를, 하나는 최강력의 정의를 거친다.

B3 힌트 2회 정리 1.13은 꼬리 공식 $EN = \sum_{i \geq 1} P(N \geq i)$이다. 조건 확인에는 $|Z_i|$에 같은 계산을 한 결과를 쓴다.

B4 힌트 B1에서 채운 근거를 다시 보지 말고 쓴다. (a)의 $F(k-)$는 $P_{\theta_0}(f_1 \geq kf_0)$이다.

B5 힌트 "위쪽 집합"은 $\{f'' > kf'\} \subset \{\varphi^* = 1\}$, $\{f'' < kf'\} \subset \{\varphi^* = 0\}$이라는 뜻이다.

B6 힌트 (b)의 부등식에서 $k \geq 0$을 어디에 쓰는지 표시한다. (c)에서는 φ의 크기가 α 이하라는 것도 쓴다.

B7 힌트 노트 연습 4.1의 반례에서 $\{N \geq 2\}$는 무엇으로 정해지는가.

C1 힌트 순증가함수는 부등호의 방향을 보존하고, 대우로 거꾸로도 보존한다. $f_0 = 0$인 점은 H_0 아래 확률 0이다.

C2 힌트 (a) A 위에서는 가능도비가 1보다 크고 B 위에서는 작다. 가능도비가 T의 감소하지 않는 함수 g이면, $g(T(\mathbf{x})) > g(T(\mathbf{y}))$에서 $T(\mathbf{x}) > T(\mathbf{y})$가 나온다(대우).

C3 힌트 $\theta < 1$이면 가능도비 f_θ/f_1은 $\{X_{(n)} \leq \theta\}$ 위에서 상수 θ^{-n}이고 그 밖에서 0이다. 그래서

어떤 검정이든 검정력 $= \theta^{-n} \times$(그 집합 위에서 쓴 크기)이다.

답 $\dfrac{0.05}{0.8^5} \approx 0.153$

C4 힌트 H_0 아래 $U = (X_{(n)}/\theta_0)^n \sim \mathrm{U}(0,1)$이고, $-2\log U$는 평균 2인 지수분포, 곧 $\chi^2(2)$이다.

답 $e^{-3.841/2} \approx 0.146$

C5 힌트 (a) $S_N = \sum_{i \le m} Z_i 1\{N \ge i\}$를 제곱한다. $i < j$이면 $Z_i 1\{N \ge i\}$는 $Z_1, \dots, Z_{j-1}$로 정해지므로 Z_j와 독립이다.

답 $EN = ab = 15$

C6 힌트 귀류법 [T2]. 지배하는 δ'가 있다고 하고 Bayes 위험 $\sum_j \pi_j R(\theta_j, \cdot)$을 비교한다.

C7 힌트 (a)의 첫 부등호: 평균은 최댓값을 넘지 못한다. 둘째: Bayes 규칙의 정의.

C8 힌트 (b) 크기 0.05 이하인 부분집합은 몇 개 안 된다. (c) 가능도비가 큰 것부터 크기를 채우고 남은 크기만큼 다음 결과에서 확률화한다.

답 (b) 0.089 (c) 0.098

C9 힌트 (a) 두 검정을 t와 $1-t$로 섞은 $t\varphi_1 + (1-t)\varphi_2$도 검정이다. (c) 오목성: 크기 α_1, α_2의 최강력 검정을 섞은 검정의 크기와 검정력을 보고 γ^*의 정의와 비교한다.

D1 힌트 (a) 결합밀도는 $e^{n\theta} e^{-\sum x_i} 1\{x_{(1)} > \theta\}$이다. $\theta_1 < \theta_2$의 비는 $x_{(1)} \le \theta_2$이면 0, 그 위에서는 상수다. (b) $P(X_{(1)} > t) = P(X_1 > t)^n$.

답 (a) $c = \dfrac{\log 20}{10} \approx 0.300$ (b) $\dfrac{e^2}{20} \approx 0.369$

D2 힌트 (a) H_0 아래 MLE는 $\frac{x+y}{2}$, 전체 모형에서는 (x, y)이다. (b) H_0 아래 $X+Y$는 공통 λ의 충분통계량이다. $P(X = k \mid X+Y = n)$을 직접 계산한다.

답 (a) ≈ 4.19 (b) $\dfrac{2517}{65536} \approx 0.0384$

D3 힌트 불량품마다 Λ에 $\log 3$을, 양품마다 $\log \frac{7}{9}$를 더한다. 그래서 $\Lambda_n = d_n \log \frac{27}{7} - n \log \frac{9}{7}$이고, $A = -B = \log 19$이다.

답 (a) $s \approx 0.186$ (b) $h \approx 2.18$ (c) $E_0 N \approx 22.8$ (d) $E_1 N \approx 17.2$ (e) 12 (f) 3

D4 힌트 (a) 가능도비 e^{2x-2}를 임계값 $\frac{0.8 \cdot 1}{0.2 \cdot 5} = 0.8$과 비교한다. (c) $1 - \Phi(t) = 5\Phi(t-2)$를 t에 대해 이분법으로 푼다.

답 (a) $t \approx 0.888$ (b) ≈ 0.283 (c) $t \approx 0.474$