수리통계학 × 통계학의 역사

7회 워크북

극한의 언어
손으로 푸는 34문항

강의노트 7회를 읽은 뒤 손으로 푼다. 극한 정리는 가정을 하나씩 확인하는 일이 증명의 대부분이다. 어느 정리를 어떤 가정 아래 쓰는지 줄마다 적는다. 문제 아래의 줄 친 공간에 직접 쓴다. 정규분포의 값이 필요한 문항은 문제에 그 값을 적어 두었다.

부	하는 일	문항	예상 시간
A 계산 드릴	정규근사, 델타 방법, Monte Carlo의 분산을 손으로 계산한다	14	1시간 50분
B 빈칸 증명	노트의 단계형 증명을 책을 덮고 다시 쓴다	7	2시간 50분
C 증명 문제	변형, 일반화, 반례 만들기	9	2시간 50분
D 퀄형	여러 절을 잇는 종합 문제. 제한 시간 안에	4	2시간 25분
	합계	34	약 10시간

노트 `07_limits_notes.pdf` — 문항의 "정리 3.1" 같은 번호는 노트의 번호다

도구 `00_toolkit_notes.pdf` — [T1] 원소 추적, [T2] 반례, [T4] 부등식, [T5] 극한과 급수, [T8] Taylor 전개

검산 노트북 `07_limits.ipynb`의 "워크북 검산" 절: `check("A5", "1/4")`

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. 정규분포표의 값으로 계산한 답은 넷째 자리까지 맞으면 맞다고 본다.
5. **극한 정리는 가정을 적는다.** 대수의 법칙, 중심극한정리, Slutsky, 델타 방법을 쓸 때마다 그 정리의 가정(평균의 존재, 분산의 존재, 상수로의 확률수렴, 미분가능성)을 확인한 줄을 남긴다.

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. 필요한 정규분포의 값은 문제에 적어 두었다.

A1 Chebyshev와 중심극한정리 ★

정리 3.1, 2회 정리 5.3 · 8분

성공 확률 0.3인 독립 시행 n번의 성공 비율을 $\hat{p}$라 하자. $P(|\hat{p}-0.3|\ge 0.02)\le 0.05$가 되는 가장 작은 n을 (a) Chebyshev 부등식으로, (b) 중심극한정리의 근사로($z_{0.025}=1.96$) 구하라.

A2 이항의 정규근사 ★

따름정리 3.2(a) · 8분

$K\sim \mathrm{b}(400,\tfrac{1}{2})$에서 $P(K\le 190)$을 (a) 연속성 보정을 하여, (b) 보정 없이 정규근사로 구하라. $\Phi(0.95)=0.8289$, $\Phi(1)=0.8413$이다. 정확한 값은 0.1711이다.

A3 Poisson의 정규근사 ★

따름정리 3.2(b) · 6분

$X \sim \mathrm{Poi}(100)$에서 $P(X \geq 115)$를 연속성 보정을 한 정규근사로 구하라. $\Phi(1.45) = 0.9265$이다. 정확한 값 0.0759와 비교하라.

A4 최댓값의 극한 ★

예제 2.2(b) · 6분

$U_1, \ldots, U_{10} \overset{\text{iid}}{\sim} \mathrm{U}(0,1)$의 최댓값을 M_{10}이라 하자. (a) $P\big(10(1 - M_{10}) > 2\big)$을 정확히 구하라. (b) 극한분포 $\mathrm{Exp}(1)$로 근사한 값을 구하라.

A5 Poisson 평균의 제곱근 ★

정리 4.4 · 6분

$X_1, \ldots, X_n \overset{\text{iid}}{\sim} \mathrm{Poi}(\lambda)$이다. $\sqrt{n}\big(\sqrt{\bar{X}_n} - \sqrt{\lambda}\big)$의 극한분포의 분산을 구하라.

A6 지수분포의 비율 모수 ★

정리 4.4 · 8분

$X_1, \ldots, X_n$이 iid이고 밀도가 $\lambda e^{-\lambda x}$ $(x > 0)$이다. $\hat{\lambda} = 1/\bar{X}_n$에 대해 $\sqrt{n}(\hat{\lambda} - \lambda)$의 극한분포의 분산을 λ의 식으로 구하라. (노트북의 `check()`에는 λ를 `lam`으로 쓴다.)

A7 Monte Carlo의 표본 크기 ★

정리 6.1 · 10분

$\theta = \int_0^1 \sqrt{1-x^2}\,dx = \frac{\pi}{4}$를 $U_i \overset{\text{iid}}{\sim} \mathrm{U}(0,1)$과 $h(u) = \sqrt{1-u^2}$의 평균으로 추정한다. (a) $\sigma_h^2 = \mathrm{Var}\, h(U)$를 정확히 구하라. (b) 표준오차 $\sigma_h/\sqrt{n}$이 0.001 이하가 되는 가장 작은 n을 구하라.

A8 베타 난수의 기각 표집 ★

알고리즘 6.3, 정리 6.4 · 8분

$f(x) = 6x(1-x)$ $(0 < x < 1)$를 제안분포 $\mathrm{U}(0,1)$로 기각 표집한다. (a) 쓸 수 있는 가장 작은 M은? (b) 한 번 시도에서 받아들일 확률은? (c) 정확히 세 번째 시도에서 처음 받아들일 확률은?

A9 평균의 제곱 ★

정리 4.4 · 6분

$X_i \overset{\text{iid}}{\sim} \mathrm{Exp}(1)$이면 $\sqrt{n}(\bar{X}_n^2 - 1)$의 극한분포의 분산은 얼마인가?

A10 로그 오즈의 분산 ★

예제 4.5 · 6분

$p = 0.2$, $n = 100$일 때 $\log\frac{\hat{p}}{1-\hat{p}}$의 근사 분산을 델타 방법으로 구하라.

A11 세 범주의 비율 ★★

예제 5.3 · 10분

세 범주의 확률이 $\mathbf{p} = (0.5, 0.3, 0.2)$이고 n번의 독립 시행에서 비율 $\hat{p}_1, \hat{p}_2, \hat{p}_3$을 얻는다. (a) $\sqrt{n}(\hat{p}_1 - \hat{p}_2 - 0.2)$의 극한분포의 분산을 구하라. (b) $\sqrt{n}\hat{p}_1$과 $\sqrt{n}\hat{p}_2$의 극한 공분산을 구하라.

A12 지수 꼬리의 중요도 표집 ★★

정리 6.6 · 12분

$X \sim \text{Exp}(1)$에서 $\theta = P(X > 3) = e^{-3}$을 추정한다. 제안분포는 평균 3인 지수분포 $g(y) = \frac{1}{3}e^{-y/3}$이고 $h = 1\{y > 3\}$이다. (a) 가중치 $w = f/g$를 쓰고 $E_g[h^2 w^2]$을 구하라. (b) 한 표본의 분산의 비 $\frac{\theta(1-\theta)}{E_g[h^2w^2]-\theta^2}$를 구하라(소수 둘째 자리까지).

A13 Berry–Esseen 한계 ★

3.3절의 상자 · 6분

공정한 동전(Bernoulli($\frac{1}{2}$)) $n = 100$번의 표준화한 합에 대해 Berry–Esseen 상한 $C\rho/(\sigma^3\sqrt{n})$을 $C = 0.4748$로 구하라.

A14 반대 변량의 이득 ★★

예제 6.8 · 10분

$h(u) = u^2$, $U \sim \mathrm{U}(0,1)$이다. (a) $\mathrm{Var}\, h(U)$, (b) $\mathrm{Cov}\big(h(U), h(1-U)\big)$, (c) $\mathrm{Var}\, \frac{h(U)+h(1-U)}{2}$를 구하라. 독립인 난수 두 개의 평균과 비교하면 분산이 몇 배 줄었나?

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 앞 회차의 정리 번호, 이 회차의 정리 번호, 도구 번호를 쓴다.

B1 중심극한정리, 근거 채우기 ★★

① · 정리 3.1 · 20분

iid이고 mgf가 $|t| < h$에서 존재하며 평균 μ, 분산 $\sigma^2 > 0$이면 $Z_n = \sqrt{n}(\bar{X}_n - \mu)/\sigma \xrightarrow{D} \mathcal{N}(0,1)$이다. 각 단계의 근거를 써라.

① $Y_i = (X_i - \mu)/\sigma$는 iid, 평균 0, 분산 1이고 mgf가 $|s| < \sigma h$에서 존재한다.
왜? _____

② $\psi = \log M_Y$에 대해 $\psi(0) = 0$, $\psi'(0) = 0$, $\psi''(0) = 1$이다. 왜? _____

③ $\log M_{Z_n}(t) = n\,\psi(t/\sqrt{n})$이다. 왜? _____

④ $\psi(s) = \frac{1}{2}\psi''(\xi_n)s^2$인 ξ_n이 0과 $s = t/\sqrt{n}$ 사이에 있다. 왜? _____

⑤ $n\psi(t/\sqrt{n}) = \frac{t^2}{2}\psi''(\xi_n) \to \frac{t^2}{2}$이다. 왜? _____

⑥ $M_{Z_n}(t) \to e^{t^2/2}$이므로 $Z_n \xrightarrow{D} \mathcal{N}(0,1)$이다. 왜? _____

B2 확률수렴은 분포수렴, 근거 채우기 ★★

① · 정리 2.3(a) · 15분

$X_n \xrightarrow{P} X$이면 $X_n \xrightarrow{D} X$이다. 근거를 써라.

① $\{X_n \le x\} \subset \{X \le x+\varepsilon\} \cup \{|X_n - X| \ge \varepsilon\}$이다. 왜? _____

② $\{X \le x-\varepsilon\} \subset \{X_n \le x\} \cup \{|X_n - X| \ge \varepsilon\}$이다. 왜? _____

③ $F(x-\varepsilon) - P(|X_n - X| \ge \varepsilon) \le F_n(x) \le F(x+\varepsilon) + P(|X_n - X| \ge \varepsilon)$이다.

왜? _____

④ $F(x-\varepsilon) \le \liminf_n F_n(x) \le \limsup_n F_n(x) \le F(x+\varepsilon)$이다.

왜? _____

⑤ x가 F의 연속점이면 $\lim_n F_n(x) = F(x)$이다. 왜? _____

B3 델타 방법, 근거 채우기 ★★

① · 정리 4.4 · 15분

$\sqrt{n}(Y_n - \theta) \xrightarrow{D} \mathcal{N}(0, \sigma^2)$이고 g가 θ에서 미분가능하면 $\sqrt{n}(g(Y_n) - g(\theta)) \xrightarrow{D} \mathcal{N}(0, g'(\theta)^2\sigma^2)$이다. 근거를 써라.

① $h(y) = \frac{g(y)-g(\theta)}{y-\theta}$ $(y \ne \theta)$, $h(\theta) = g'(\theta)$는 θ에서 연속이다. 왜? _____

② $Y_n - \theta = \frac{1}{\sqrt{n}} \cdot \sqrt{n}(Y_n - \theta) \xrightarrow{P} 0$이다. 왜? _____

③ $h(Y_n) \xrightarrow{P} g'(\theta)$이다. 왜? _____

④ $\sqrt{n}(g(Y_n) - g(\theta)) = h(Y_n) \cdot \sqrt{n}(Y_n - \theta)$이다. 왜? _____

⑤ 극한은 $g'(\theta) \cdot \mathcal{N}(0, \sigma^2) = \mathcal{N}(0, g'(\theta)^2\sigma^2)$이다. 왜? _____

B4 중심극한정리, 단계의 이름만 ★★

② · 정리 3.1 · 30분

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

① **표준화한다**

② **로그 mgf의 0에서의 값과 두 도함수를 적는다**

③ **독립합의 mgf로 $\log M_{Z_n}$ 을 쓴다**

④ **나머지항이 있는 Taylor 전개를 한다**

⑤ **n을 곱한 것의 극한을 구한다**

⑥ **mgf의 극한에서 분포수렴으로 간다**

B5 두 수렴의 관계, 단계의 이름만 ★★

② · 정리 2.3 · 25분

정리 2.3의 (a)와 (b)를 증명하라.

① **(a) 두 포함관계를 쓴다**

② **(a) 확률을 취해 $F_n(x)$를 끼운다**

③ **(a) 상극한·하극한을 취하고 $\varepsilon \downarrow 0$**

④ **(b) 상수의 분포함수와 그 연속점을 적는다**

⑤ **(b) $\{|X_n - c| \geq \varepsilon\}$을 분포함수로 덮는다**

⑥ **(b) 극한을 취한다**

B6 중심극한정리, 전략만 ★★★

③ · 정리 3.1 · 35분

정리 3.1을 증명하라. 정의역에 관한 조건 확인도 함께 쓴다.

전략 표준화한 합의 mgf의 로그를 $n\psi(t/\sqrt{n})$으로 쓰고, ψ를 0 근처에서 2차까지 전개한다. 2차 항만 n을 곱해도 남는다.

B7 델타 방법, 전략만 ★★

③ · 정리 4.4 · 30분

정리 4.4를 증명하라. 그리고 $g(y)=|y|$, $\theta=0$이면 증명의 어느 단계가 무너지는지 밝혀라.

전략 차분몫 $h(y)=\frac{g(y)-g(\theta)}{y-\theta}$를 확률변수 계수로 보고, 그 계수가 $g'(\theta)$로 확률수렴함을 보인 뒤 Slutsky를 쓴다.

C

증명 문제

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

C1 극한이 확률변수인 사칙연산 ★★

[T1], [T4] · 정리 1.3, 2.5 · 15분

$X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y$이다(X, Y는 확률변수). (a) $X_n + Y_n \xrightarrow{P} X + Y$임을 보여라. (b) $X_n Y_n \xrightarrow{P} XY$임을 보여라. (힌트: $X_n Y_n - XY = (X_n - X)Y_n + X(Y_n - Y)$.)

C2 평균제곱수렴과 확률수렴 ★★

[T2] · 2회 따름정리 5.2 · 15분

(a) $E(Y_n - c)^2 \to 0$이면 $Y_n \xrightarrow{P} c$임을 보여라. (b) $P(Y_n = \sqrt{n}) = \frac{1}{n}$, $P(Y_n = 0) = 1 - \frac{1}{n}$이면 $Y_n \xrightarrow{P} 0$이고 $EY_n \to 0$이지만 EY_n^2은 0으로 가지 않음을 보여라. 따라서 (a)의 역은 거짓이다.

C3 유계이면 기댓값도 따라온다 ★★

[T4] · 정의 1.2 · 15분

$Y_n \xrightarrow{P} c$이고 모든 n에서 $|Y_n| \le K$이면 $EY_n \to c$임을 보여라. 유계라는 가정이 없으면 왜 성립하지 않는지 노트 연습 1.3의 예로 설명하라.

C4 독립인 두 수열의 합 ★★

2회 정리 3.6 · 2.3절의 상자 · 20분

각 n에서 X_n과 Y_n이 독립이고, mgf가 0 근방에서 $M_{X_n} \to M_X$, $M_{Y_n} \to M_Y$로 수렴한다(M_X, M_Y는 X, Y의 mgf). (a) $X_n + Y_n \xrightarrow{D} X' + Y'$임을 보여라. 여기서 X', Y'은 X, Y와 각각 같은 분포이고 서로 독립이다. (b) 노트 연습 2.5의 $Y_n = -X_n$이 (a)의 반례가 아닌 이유는?

C5 미분가능하지 않으면 ★★

[T2] · 정리 4.4 · 20분

$\sqrt{n}Y_n \xrightarrow{D} \mathcal{N}(0,\sigma^2)$이고 $g(y)=|y|$이다($\theta=0$). (a) $P(\sqrt{n}|Y_n|\le y)=F_n(y)-P(\sqrt{n}Y_n<-y)$로 $\sqrt{n}|Y_n| \xrightarrow{D} \sigma|Z|$임을 보여라($F_n$은 $\sqrt{n}Y_n$의 분포함수). (b) 극한의 평균 $E|Z|$을 구하라. (c) 극한이 정규분포가 아님을 보이고, 정리 4.4의 어느 가정이 깨졌는지 밝혀라.

C6 Wald 구간의 포함확률 ★★★

정리 4.2, 예제 2.2(c) · 25분

성공 확률 p인 독립 시행 n번의 비율 $\hat p$로 만든 구간 $\hat p \pm z_{\alpha/2}\sqrt{\hat p(1-\hat p)/n}$을 Wald 구간이라 한다. (a) $p\in(0,1)$이 고정이면 포함확률이 $1-\alpha$로 감을 보여라. (b) $p=p_n=1/n$이면 $P(\hat p=0)\to e^{-1}$임을 보이고, 포함확률의 상극한이 $1-e^{-1}$ 이하임을 보여라.

C7 정규화 상수를 모를 때의 중요도 표집 ★★

정리 1.3, 1.6, 6.6 · 20분

목표 밀도가 $f = \tilde{f}/c$이고 $\tilde{f}$는 알지만 상수 c는 모른다. $Y_i \overset{\text{iid}}{\sim} g$, $\tilde{w} = \tilde{f}/g$에 대해

$$\hat{\theta}_n = \frac{\sum_{i=1}^{n} h(Y_i)\tilde{w}(Y_i)}{\sum_{i=1}^{n} \tilde{w}(Y_i)}$$

로 둔다($E_f|h| < \infty$, $\tilde{f} > 0$인 곳에서 $g > 0$). (a) $\hat{\theta}_n \xrightarrow{P} E_f h(X)$임을 보여라. (b) $n = 1$이면 $\hat{\theta}_1$의 기댓값은 무엇인가? 이 추정량은 일반적으로 불편인가?

C8 독립이 아니어도 대수의 법칙 ★★

[T5] · 2회 정리 5.3 · 20분

$X_1, X_2, \dots$ 의 평균이 모두 μ이고 분산이 모두 $v < \infty$이다. (a) 쌍마다 상관이 0이면 $\bar{X}_n \xrightarrow{P} \mu$ 임을 보여라. (b) $\mathrm{Cov}(X_i, X_j) = v\rho^{|i-j|}$ ($|\rho| < 1$)이면 $\mathrm{Var}\,\bar{X}_n \le \frac{v}{n} \cdot \frac{1+|\rho|}{1-|\rho|}$ 임을 보이고 $\bar{X}_n \xrightarrow{P} \mu$ 를 결론지어라. (c) 모든 i에서 $X_i = Z$이면 무엇이 깨지는가?

C9 기각 표집의 상수와 꼬리 ★★

정리 6.4 · 20분

(a) 알고리즘 6.3이 정당하려면 $M \ge \sup_x f(x)/g(x)$이어야 하고, 받아들일 확률은 $M = \sup_x f/g$에서 가장 큼을 보여라. (b) 목표 $f = \mathcal{N}(0,1)$, 제안 $g =$ 표준 Cauchy이면 가장 작은 M이 $\sqrt{2\pi/e}$임을 보여라. (c) 목표가 Cauchy, 제안이 $\mathcal{N}(0,1)$이면 쓸 수 있는 M이 없음을 보여라.

D

퀄형

여러 절과 앞 회차의 도구를 잇는 종합 문제다. 머리의 제한 시간을 재고, 노트를 덮고 푼다. 어느 극한 정리를 어떤 가정 아래 쓰는지 밝히는 것까지 답안의 일부다.

D1 Poisson에서 0의 확률 ★★★

정리 1.6, 3.1, 4.4 · 4회 · 제한 35분

$X_1, \dots, X_n \overset{\text{iid}}{\sim} \text{Poi}(\lambda)$에서 $\theta = P(X=0) = e^{-\lambda}$를 추정한다. 두 추정량

$$T_1 = \frac{1}{n}\#\{i : X_i = 0\}, \qquad T_2 = e^{-\bar{X}_n}$$

을 비교한다. (노트북의 `check()`에는 λ를 `lam`으로 쓴다.)

(a) T_1과 T_2가 모두 θ로 확률수렴함을 보여라.

(b) $\sqrt{n}(T_1 - \theta)$의 극한분포의 분산을 구하라.

(c) $\sqrt{n}(T_2 - \theta)$의 극한분포의 분산을 구하라.

(d) 점근분산의 비 $\frac{\text{(c)}}{\text{(b)}}$가 모든 $\lambda > 0$에서 1보다 작음을 보이고, $\lambda = 1$에서의 값을 구하라.

D2 균등분포의 끝점 ★★★

예제 2.2(b), 정리 3.1 · 3회 · 제한 40분

$X_1, \dots, X_n \overset{\text{iid}}{\sim} \mathrm{U}(0,\theta)$이고 $M_n = \max_i X_i$이다.

(a) $x \ge 0$에서 $P\big(n(\theta - M_n) > x\big) \to e^{-x/\theta}$ 임을 보여라.

(b) $2\bar{X}_n$의 분산을 구하라.

(c) $T = \frac{n+1}{n} M_n$이 불편이고 분산이 $\frac{\theta^2}{n(n+2)}$ 임을 보여라.

(d) $n = 10$에서 $\mathrm{Var}(2\bar{X}_n)/\mathrm{Var}\,T$를 구하라. 두 추정량의 오차는 n에 대해 각각 어떤 속도로 줄어드는가?

D3 기하분포의 성공 확률 ★★★

정리 3.1, 4.2, 4.4 · 4회 · 제한 30분

$X_1, \dots, X_n$이 iid이고 $P(X=k) = (1-p)^{k-1}p$ $(k = 1, 2, \dots)$이다. $EX = \frac{1}{p}$, $\operatorname{Var} X = \frac{1-p}{p^2}$이다.

(a) $\hat{p} = 1/\bar{X}_n$에 대해 $\sqrt{n}(\hat{p} - p)$의 극한분포의 분산을 구하라.

(b) 그 분산의 p 자리에 $\hat{p}$를 넣은 근사 95% 구간이 점근적으로 옳음을 Slutsky로 보여라.

(c) $n = 50$, $\bar{X} = 4$이면 (b)에서 쓰는 표준오차 $\hat{p}\sqrt{(1-\hat{p})/n}$은 얼마인가?

(d) 같은 자료로 (b)의 구간을 구하라($z_{0.025} = 1.96$).

D4 정규 꼬리의 중요도 표집 ★★★

정리 6.6 · 5회 · 제한 40분

$\theta = P(Z > a) = 1 - \Phi(a)$를 제안분포 $g = \mathcal{N}(\mu, 1)$의 중요도 표집으로 추정한다. $h = 1\{y > a\}$, $w(y) = \phi(y)/\phi(y-\mu)$이다.

(a) $w(y) = e^{-\mu y + \mu^2/2}$이고 $E_g[h^2 w^2] = e^{\mu^2}\big(1 - \Phi(a+\mu)\big)$임을 보여라.

(b) $a = \mu = 3$일 때 한 표본의 분산의 비 $\dfrac{\theta(1-\theta)}{E_g[h^2w^2] - \theta^2}$을 구하라. $1 - \Phi(3) = 1.3499 \times 10^{-3}$, $1 - \Phi(6) = 9.866 \times 10^{-10}$, $e^9 = 8103.08$이다.

(c) $E_g[h^2w^2]$을 최소로 하는 μ가 $2\mu\big(1-\Phi(a+\mu)\big) = \phi(a+\mu)$를 만족함을 보이고, $1-\Phi(x) \approx \phi(x)/x$를 써서 $\mu \approx a$임을 보여라.

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 $\operatorname{Var}\hat{p} = pq/n$이다. (a)는 $\frac{pq}{n\varepsilon^2} \le 0.05$, (b)는 $z_{0.025}\sqrt{pq/n} \le 0.02$를 푼다.

답 (a) 10,500 (b) 2,017

A2 힌트 $EK = 200$, $\operatorname{sd}(K) = 10$이다. 연속성 보정은 $K \le 190$을 $K \le 190.5$로 읽는다.

답 (a) $\Phi(-0.95) = 0.1711$ (b) $\Phi(-1) = 0.1587$

A3 힌트 $EX = \operatorname{Var}X = 100$이다. $X \ge 115$는 $X \ge 114.5$로 읽는다.

답 $1 - \Phi(1.45) = 0.0735$

A4 힌트 $10(1 - M_{10}) > 2$는 $M_{10} < 0.8$, 곧 열 개가 모두 0.8보다 작다는 것이다.

답 (a) $0.8^{10} = 0.1074$ (b) $e^{-2} = 0.1353$

A5 힌트 중심극한정리로 $\sqrt{n}(\bar{X}_n - \lambda) \xrightarrow{D} \mathcal{N}(0, \lambda)$이다. $g(x) = \sqrt{x}$에 델타 방법.

답 $\frac{1}{4}$

A6 힌트 $EX = 1/\lambda$, $\operatorname{Var}X = 1/\lambda^2$이다. $g(x) = 1/x$를 $1/\lambda$에서 미분한다.

답 λ^2

A7 힌트 $Eh(U)^2 = \int_0^1 (1 - u^2)\,du$이다. $\pi^2/16 = 0.61685$이다.

답 (a) $\frac{2}{3} - \frac{\pi^2}{16} \approx 0.04982$ (b) 49,817

A8 힌트 $M = \sup_x f(x)/g(x)$이다. 시도 횟수는 성공 확률 $1/M$의 기하분포다.

답 (a) $\frac{3}{2}$ (b) $\frac{2}{3}$ (c) $\frac{2}{27}$

A9 힌트 $\mu = \sigma^2 = 1$이다. $g(x) = x^2$의 $g'(1)$.

답 4

A10 힌트 점근분산 $\frac{1}{pq}$를 n으로 나눈다.

답 $\frac{1}{16} = 0.0625$

A11 힌트 극한 공분산행렬은 $\mathrm{diag}(\mathbf{p}) - \mathbf{p}\mathbf{p}'$이다. (a)는 $\mathbf{a} = (1, -1, 0)$에 대한 $\mathbf{a}'\Sigma\mathbf{a}$.

답 (a) 0.76 (b) -0.15

A12 힌트 $w(y) = 3e^{-2y/3}$이다. $E_g[h^2 w^2] = \int_3^\infty 9e^{-4y/3} \cdot \frac{1}{3}e^{-y/3}\,dy$.

답 (a) $\frac{9}{5}e^{-5} \approx 0.01213$ (b) ≈ 4.90

A13 힌트 $|X - \frac{1}{2}|$은 언제나 $\frac{1}{2}$이다.

답 0.04748

A14 힌트 $EU^2 = \frac{1}{3}$, $EU^4 = \frac{1}{5}$, $E[U^2(1-U)^2] = \int_0^1 u^2(1-u)^2 du = \frac{1}{30}$.

답 (a) $\frac{4}{45}$ (b) $-\frac{7}{90}$ (c) $\frac{1}{180}$

B1 힌트 mgf에 관한 근거는 모두 2회에 있다. ⑤에서 $|\xi_n| \le |t|/\sqrt{n}$이다.

B2 힌트 ③은 포함관계에 단조성과 Boole 부등식을 쓴 것이다.

B3 힌트 ②는 분포수렴하는 수열이 확률적으로 유계라는 것에서 나온다.

B4 힌트 B1에서 채운 근거를 다시 보지 말고 쓴다. 넷째 단계에서 1차 항이 없어지는 이유를 적는다.

B5 힌트 (b)의 덮기에서 $c + \varepsilon$ 대신 $c + \frac{\varepsilon}{2}$를 쓰는 이유를 생각한다. $P(X_n \ge c + \varepsilon)$은 분포함수의 값이 아니다.

B6 힌트 단계는 여섯이다. 조건 확인에서는 $t/\sqrt{n}$이 ψ의 정의역 안에 있는지, $\log M_Y$가 정의되는지, mgf의 수렴이 0을 포함하는 열린 구간에서 성립하는지를 본다.

B7 힌트 $g(y) = |y|$이면 $h(y) = \mathrm{sign}(y)$이다. 이 h는 0에서 연속인가?

C1 힌트 (a) $|X_n + Y_n - X - Y| \ge \varepsilon$이면 두 차 가운데 하나는 $\varepsilon/2$ 이상이다. (b) 확률수렴하면 분포수렴하고(정리 2.3) 그러면 확률적으로 유계다(정리 2.5). 확률변수 하나 X도 확률적으로 유계다.

C2 힌트 (a) 일반화된 Markov 부등식 $P(|Y_n - c| \ge \varepsilon) \le E(Y_n - c)^2/\varepsilon^2$.

C3 힌트 $E|Y_n - c|$을 $\{|Y_n - c| < \varepsilon\}$ 위와 그 바깥으로 나눈다. 바깥에서는 $|Y_n - c| \le K + |c|$이다.

C4 힌트 독립합의 mgf는 mgf의 곱이다.

C5 힌트 (a) 노트 연습 2.4의 끼우기와 같다. (b) $E|Z| = 2\int_0^\infty z\phi(z)\,dz$.

답 (b) $\sqrt{2/\pi} \approx 0.798$

C6 힌트 (a) $\dfrac{\hat p - p}{\sqrt{\hat p(1-\hat p)/n}} = \sqrt{\dfrac{pq}{\hat p(1-\hat p)}} \cdot \dfrac{\hat p - p}{\sqrt{pq/n}}$. (b) $\hat p = 0$이면 구간은 한 점 $\{0\}$이다.

답 (b) $1 - e^{-1} \approx 0.632$

C7 힌트 분자와 분모를 각각 n으로 나누면 iid 평균이다. $E_g \tilde w = c$이다.

C8 힌트 $\operatorname{Var}\bar X_n = \frac{1}{n^2}\sum_{i,j}\operatorname{Cov}(X_i, X_j)$. (b)에서 고정된 i에 대해 $\sum_j |\rho|^{|i-j|} \le 1 + 2\sum_{k\ge 1}|\rho|^k$.

C9 힌트 (b) $f/g = \sqrt{\pi/2}\,(1+x^2)e^{-x^2/2}$을 미분한다. (c) $x \to \infty$에서 f/g를 본다.

답 (b) $\sqrt{2\pi/e} \approx 1.520$

D1 힌트 (b) $1\{X_i = 0\}$은 Bernoulli(θ)다. (c) $g(x) = e^{-x}$에 델타 방법. (d) 비는 $\frac{\lambda}{e^\lambda - 1}$이고 $e^\lambda > 1 + \lambda$ [T4].

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

D2 힌트 (a) $P(M_n < \theta - x/n) = (1 - \frac{x}{n\theta})^n$. (c) M_n/θ의 밀도는 nu^{n-1} $(0<u<1)$이다.

답 (b) $\dfrac{\theta^2}{3n}$ (c) $\dfrac{\theta^2}{n(n+2)}$ (d) 4

D3 힌트 (a) $g(x) = 1/x$, $g'(1/p) = -p^2$. (c) $\hat p = 0.25$이다. (d) 노트북의 `check()`에는 구간의 아래 끝을 넣는다.

답 (a) $p^2(1-p)$ (c) 0.0306 (d) $[0.190,\ 0.310]$

D4 힌트 (a) $\phi(y-\mu)e^{-2\mu y + \mu^2}$의 지수를 완전제곱으로 정리하면 $\phi(y+\mu)e^{\mu^2}$이다.

답 (b) ≈ 218