수리통계학 × 통계학의 역사

4회 워크북

셈하는 분포와 섞인 분포 손으로 푸는 35문항

강의노트 4회를 읽은 뒤 손으로 푼다. 분포의 pmf와 밀도를 손으로 합하고 적분하며, 지시변수, 변수변환, 혼합의 세 도구를 빈 종이에서 다시 쓴다. 문제 아래의 줄 친 공간에 직접 쓴다.

부	하는 일	문항	예상 시간
A 계산 드릴	pmf · 밀도 · 적률 · 사후분포를 손으로 계산한다	14	2시간 10분
B 빈칸 증명	노트의 단계형 증명을 책을 덮고 다시 쓴다	7	2시간 35분
C 증명 문제	변형, 일반화, 반례 만들기	10	3시간 25분
D 퀄형	여러 회차를 잇는 종합 문제. 제한 시간 안에	4	2시간 35분
	합계	35	약 10시간 45분

노트 `04_counting_notes.pdf` — 문항의 "정리 3.3" 같은 번호는 노트의 번호다. 앞 회차는 "3회 정리 5.1" 처럼 쓴다

도구 `00_toolkit_notes.pdf` — [T1] 지시함수, [T2] 귀류, [T3] 귀납, [T4] 부등식, [T5] 극한, [T6] 합과 적분의 교환, [T7] 변수변환, [T8] 급수 전개

검산 노트북 `04_counting.ipynb`의 "워크북 검산" 절: `check("A10c", "11/16")`

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.** 초기하분포의 분산, 무기억성의 역, 베타와 감마의 관계를 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.
5. **증명은 근거를 단다.** 등호와 부등호마다 어느 정의, 정리, 도구를 썼는지 오른쪽에 적는다. 합과 적분의 순서를 바꿀 때는 조건을 확인한다.

표시 ★ 정의를 바로 쓰는 문제, ★★ 두세 단계를 잇는 문제, ★★★ 퀄에 나오는 길이의 문제. [T1]–[T8]은 증명 도구 상자의 도구 번호다.

예: 빈칸과 풀이 공간

① $X \sim \mathrm{Poi}(\lambda)$이면 $EX = \lambda$이다. 왜? _____

A

계산 드릴

pmf, 밀도, 적률, 사후분포를 손으로 계산한다. 답이 수나 식으로 떨어지므로 다 풀면 뒤의 최종 답이나 노트북의 `check()`로 맞춰 본다. 계산기는 마지막 수치에만 쓴다.

A1 세 번째 6까지 ★

정리 1.3 · 8분

공정한 주사위를 6이 세 번 나올 때까지 던진다. (a) 6이 아닌 눈이 정확히 5번 나올 확률을 구하라. (b) 던지는 횟수의 평균을 구하라.

A2 브리지 패의 하트 ★

정의 1.4, 정리 1.5 · 10분

잘 섞은 52장에서 13장을 받는다. 하트의 수를 X라 하자. (a) EX, (b) $\operatorname{Var} X$, (c) $P(X = 0)$을 구하라.

A3 열 명의 혈액형 ★★

정의 1.7, 정리 1.8 · 10분

혈액형 A, B, O, AB의 비율이 0.34, 0.27, 0.28, 0.11인 집단에서 열 명을 독립적으로 뽑는다. (a) A, B, O가 3명씩이고 AB가 1명일 확률을 구하라. (b) A형의 수와 B형의 수의 공분산을 구하라. (c) 두 수의 상관계수를 구하라.

A4 불량품 셋 이상 ★

정리 2.3 · 10분

불량률이 0.002인 공정에서 1000개를 만든다(각각 독립). 불량품이 3개 이상일 확률을 (a) Poisson 근사로, (b) 이항분포로 정확히 구하라. (c) 2.1절의 Le Cam 한계 np^2은 얼마인가? (a)와 (b)의 차이와 비교하라.

A5 콜센터 ★

정의 3.1, 정리 3.2 · 10분

전화가 시간당 평균 6통씩 Poisson 과정으로 걸려 온다. 시간은 분 단위로 쓴다. (a) 10분 동안 전화가 한 통도 없을 확률을 구하라. (b) 세 번째 전화가 올 때까지의 시간의 평균과 (c) 분산을 구하라. (d) 첫 전화가 10분과 20분 사이에 올 확률을 구하라.

A6 방문자와 구매자 ★★

정리 2.5 · 10분

하루 방문자 수가 $\text{Poi}(10)$이고, 방문자는 각자 독립적으로 확률 0.3으로 물건을 산다. (a) 구매자가 한 명도 없을 확률을 구하라. (b) 사지 않은 사람이 7명이었다는 것을 알 때 구매자 수의 조건부분산을 구하라. (c) 방문자가 10명이었다는 것을 알 때 구매자 수의 조건부분산을 구하라.

A7 감마함수로 적분하기 ★

정의 3.4, 정리 3.5 · 8분

다음을 감마함수로 계산하라. (a) $\Gamma(\frac{7}{2})$ (b) $\int_0^\infty x^4 e^{-2x}\,dx$ (c) $\int_0^\infty \sqrt{x}\,e^{-x}\,dx$

A8 감마분포의 적률과 확률 ★★

정리 3.2, 3.7 · 10분

$X \sim \Gamma(3, 2)$이다. (a) EX^2, (b) $E(1/X)$, (c) $P(X \le 2)$를 구하라.

A9 표준정규의 제곱합 ★

정의 3.8, 정리 3.9 · 8분

$Z_1, Z_2, Z_3 \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$이다. (a) $P(Z_1^2 + Z_2^2 > 4)$, (b) $\operatorname{Var}(Z_1^2 + Z_2^2 + Z_3^2)$, (c) $Ee^{(Z_1^2+Z_2^2+Z_3^2)/4}$를 구하라.

A10 베타분포 ★

정의 4.1, 따름정리 4.3 · 8분

$U \sim \text{Beta}(3,2)$이다. (a) EU, (b) $\operatorname{Var} U$, (c) $P(U > \frac{1}{2})$를 구하라.

A11 사후분포와 다음 시행 ★★

정리 4.5, 따름정리 4.3 · 10분

성공 확률 p의 사전분포가 $\text{Beta}(2,2)$이고 10번 가운데 7번 성공했다. (a) 다음 시행이 성공할 확률, (b) 사후분산, (c) 다음 두 시행이 모두 성공할 확률을 구하라.

A12 두 Poisson의 혼합 ★★

정의 5.1, 정리 5.2 · 10분

확률 0.3으로 Poi(1), 확률 0.7로 Poi(4)에서 뽑은 X가 있다. (a) EX, (b) $\operatorname{Var} X$를 구하라. (c) $X=0$이었을 때 Poi(1)에서 뽑았을 확률을 구하라.

A13 다섯 점의 중앙값 ★★

예제 4.4, 따름정리 4.3 · 10분

$U_1,\dots,U_5 \overset{\text{iid}}{\sim} \mathrm{U}(0,1)$의 중앙값을 M이라 하자. (a) $P(M \le \frac{1}{4})$를 구하라. (b) $\operatorname{Var} M$을 구하라.

A14 닳는 전구와 닳지 않는 전구 ★

정리 3.3 · 8분

이미 500시간 켜진 전구가 앞으로 1000시간 이상 버틸 확률을 구하라. 수명이 (a) 평균 1000시간인 지수분포일 때, (b) $\Gamma(2, 500)$일 때. 두 분포의 평균은 같다.

B

빈칸 증명

노트의 단계형 증명을 책을 덮고 다시 쓴다. 같은 정리를 여러 번 만나며 도움이 줄어든다. ① 단계는 주어지고 근거와 몇 개의 식만 빈칸이다. ② 단계의 이름만 주어진다. ③ 전략만 주어진다. 근거에는 정의나 정리의 번호, 앞 회차의 번호, 도구 번호를 쓴다.

B1 초기하분포의 분산, 근거 채우기 ★

① · 정리 1.5 · 15분

N개 가운데 K개가 성공인 모집단에서 n개를 되돌려 넣지 않고 뽑는다. $p = K/N$이다. 근거와 빈칸을 채워라.

① $\mathbf{1}_i$를 i번째로 뽑은 것이 성공이라는 지시변수라 하면 $X = \sum_{i=1}^{n} \mathbf{1}_i$이다.

왜? _____

② $P(\mathbf{1}_i = 1) = p$이므로 $EX = np$이다. 왜? _____

③ $i \neq j$이면 $P(\mathbf{1}_i \mathbf{1}_j = 1) =$ _____이다. 왜? _____

④ $\mathrm{Cov}(\mathbf{1}_i, \mathbf{1}_j) =$ _____이다. 왜? _____

⑤ 분산을 공분산의 합으로 쓴다.

$$\mathrm{Var}\, X = \sum_i \mathrm{Var}\, \mathbf{1}_i + \sum_{i \neq j} \mathrm{Cov}(\mathbf{1}_i, \mathbf{1}_j)$$

왜? _____

$$= np(1-p) - _____ \cdot \frac{p(1-p)}{N-1} = np(1-p)\frac{N-n}{N-1}$$

왜? _____

B2 무기억성의 역, 근거 채우기 ★★

① · 정리 3.3 · 15분

양수인 연속 확률변수 X가 모든 $s, t > 0$에서 $P(X > s + t \mid X > s) = P(X > t)$를 만족하면 지수분포임을 보인다. $G(x) = P(X > x)$이다. 근거를 채워라.

① 모든 $s > 0$에서 $G(s) > 0$이고 $G(s + t) = G(s)G(t)$이다. 왜? _____

② $c = G(1)$로 두면 자연수 n에서 $G(n) = c^n$이다. 왜? _____

③ $0 < c < 1$이다. 왜? _____

④ 양의 유리수 $q = \frac{n}{m}$에서 $G(q) = c^q$이다. 왜? _____

⑤ 실수 $x > 0$에서 $G(x) = c^x$이다. 왜? _____

⑥ $\lambda = -\log c$로 두면 $X \sim \mathrm{Exp}(\lambda)$이다. 왜? _____

B3 베타와 감마, 근거와 식 채우기 ★★

① · 정리 4.2 · 15분

$X \sim \Gamma(a, 1)$과 $Y \sim \Gamma(b, 1)$이 독립이고 $U = \frac{X}{X+Y}$, $V = X + Y$이다. 빈칸을 채워라.

① 역변환은 $x =$ _____, $y =$ _____이고 $(0,1) \times (0,\infty)$ 위의 일대일 대응이다. 왜? _____

② 역변환의 야코비안은 _____이다. 왜? _____

③ $f_{X,Y}(x,y) = \frac{x^{a-1}e^{-x}}{\Gamma(a)} \cdot \frac{y^{b-1}e^{-y}}{\Gamma(b)}$이다. 왜? _____

④ $f_{U,V}(u,v) = \frac{u^{a-1}(1-u)^{b-1}}{\Gamma(a)\Gamma(b)/\Gamma(a+b)} \cdot \frac{v^{a+b-1}e^{-v}}{\Gamma(a+b)}$이다. 왜? _____

⑤ U와 V는 독립이고 $V \sim \Gamma(a+b, 1)$이다. 왜? _____

⑥ $B(a,b) =$ _____이고 $U \sim \mathrm{Beta}(a,b)$이다. 왜? _____

B4 초기하분포의 분산, 단계의 이름만 ★★

② · 정리 1.5 · 25분

정리 1.5를 증명하라. 단계의 이름만 주어진다. 각 단계 아래에 식과 근거를 써라.

① **X를 지시변수의 합으로 쓴다**

② **한 지시변수의 기댓값과 EX를 구한다**

③ **두 지시변수가 함께 1일 확률을 구한다**

④ **공분산을 계산한다**

⑤ **분산을 분산과 공분산의 합으로 쓰고 정리한다**

B5 무기억성의 역, 단계의 이름만 ★★

② · 정리 3.3 · 25분

정리 3.3의 역, 곧 무기억성을 갖는 양수인 연속 확률변수는 지수분포라는 것을 증명하라.

① **성질을 생존함수의 함수방정식으로 바꾼다**

② **자연수에서 푼다**

③ **$c = G(1)$이 0과 1 사이에 있음을 보인다**

④ **양의 유리수로 넓힌다**

⑤ **실수로 넓힌다**

⑥ **지수분포임을 결론짓는다**

B6 베타와 감마, 전략만 ★★

③ · 정리 4.2 · 30분

정리 4.2를 증명하라. 결론에는 $U \perp V$, $U \sim \mathrm{Beta}(a,b)$, $V \sim \Gamma(a+b,1)$, $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$가 모두 들어가야 한다.

전략 $(X,Y) \mapsto (U,V)$의 변수변환으로 결합밀도를 구하고, 그것이 u만의 함수와 v만의 함수의 곱으로 인수분해됨을 보인다.

B7 초기하분포의 분산, 전략만 ★★ ③ · 정리 1.5 · 30분

정리 1.5를 증명하라. 그리고 $n = N$이면 분산이 0이 되는 이유를 식과 말로 설명하라.

전략 X를 지시변수의 합으로 쪼갠다. 지시변수들은 독립이 아니므로 분산에는 공분산 항이 남는다.

C

증명 문제

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

C1 빈 상자의 수 ★★

[T1] · 정리 1.5와 같은 방법 · 25분

공 n개를 상자 m개에 하나씩 독립적으로, 고르게 넣는다. 빈 상자의 수를 E라 하자.

(a) $EE = m\left(1-\frac{1}{m}\right)^n$ 임을 보여라.

(b) $\operatorname{Var} E = m\left(1-\frac{1}{m}\right)^n + m(m-1)\left(1-\frac{2}{m}\right)^n - m^2\left(1-\frac{1}{m}\right)^{2n}$ 임을 보여라.

(c) $m = n = 10$이면 EE는 얼마인가? (d) 그때 $\operatorname{Var} E$는 얼마인가?

C2 음이항의 Poisson 극한 ★★

[T5], [T8] · 정리 1.3, 2.3 · 20분

$\lambda > 0$을 고정하고 $Y_r \sim \mathrm{NB}\big(r,\ p_r = \frac{r}{r+\lambda}\big)$라 하자.

(a) $EY_r = \lambda$이고 $\mathrm{Var}\, Y_r = \lambda + \lambda^2/r$임을 보여라.

(b) 고정된 k에서 $P(Y_r = k) \to e^{-\lambda}\lambda^k/k!$ $(r \to \infty)$임을 보여라.

(c) 노트 5.2절의 Poisson–감마 혼합으로 이 극한을 해석하라.

C3 여러 범주로 나누기 ★★

정리 2.5, 정의 1.7 · 20분

$N \sim \mathrm{Poi}(\lambda)$이고, $N = n$이 주어지면 각 사건이 서로 독립적으로 범주 $1, \dots, m$에 확률 $p_1, \dots, p_m$으로 나뉜다. 범주별 수를 $N_1, \dots, N_m$이라 하자.

(a) $N_1, \dots, N_m$이 상호 독립이고 $N_j \sim \mathrm{Poi}(\lambda p_j)$임을 보여라.

(b) N이 $\mathrm{Poi}(\lambda)$가 아니라 정해진 n이면 (a)는 어떻게 바뀌는가?

C4 이항을 나누면 독립이 아니다 ★★

[T2] · 정리 2.5, 1.8 · 20분

정리 2.5에서 Poisson 가정이 필요함을 반례로 보인다. $N \sim \mathrm{b}(n, \theta)$이고 N이 주어지면 각 사건이 독립적으로 확률 p로 표시된다. 표시된 수를 N_1, 나머지를 $N_2 = N - N_1$이라 하자 ($q = 1 - p$).

(a) $(N_1, N_2, n - N)$이 $\mathrm{Mult}(n; \theta p, \theta q, 1 - \theta)$임을 보여라.

(b) $\mathrm{Cov}(N_1, N_2) = -n\theta^2 pq$임을 보여라.

(c) $n = 10$, $\theta = p = \frac{1}{2}$이면 $\mathrm{Cov}(N_1, N_2)$는 얼마인가?

C5 경쟁하는 지수분포 여럿 ★★

정리 3.3 · 3회 연습 3.4 · 20분

$X_1, \dots, X_k$가 독립이고 $X_i \sim \mathrm{Exp}(\lambda_i)$이다. $L = \min_i X_i$, J를 최솟값을 갖는 첨자라 하고 $\Lambda = \sum_i \lambda_i$로 둔다.

(a) $L \sim \mathrm{Exp}(\Lambda)$임을 보여라.

(b) $P(L > t,\ J = i) = \dfrac{\lambda_i}{\Lambda} e^{-\Lambda t}$임을 보이고, L과 J가 독립임을 결론지어라.

(c) $\lambda = (1, 2, 3)$이면 $P(J = 2)$는? (d) 그때 EL은?

C6 Poisson을 가리는 항등식 ★★★

[T3] · 정의 2.1 · 25분

(a) $X \sim \mathrm{Poi}(\lambda)$이면 모든 유계함수 g에 대해 $E[\lambda g(X+1)] = E[X g(X)]$임을 보여라.

(b) (a)가 $g(x) = x^2$에서도 성립한다고 받아들이고, $EX^3 = \lambda^3 + 3\lambda^2 + \lambda$임을 보여라.

(c) 거꾸로, 음이 아닌 정수값 확률변수 X가 모든 유계함수 g에 대해 (a)의 항등식을 만족하면 $X \sim \mathrm{Poi}(\lambda)$임을 보여라.

C7 섞으면 0이 늘어난다 ★★

[T4] · 정리 5.2 · 20분

$\Lambda \ge 0$이 $E\Lambda = \mu < \infty$인 확률변수이고, $\Lambda = \lambda$가 주어지면 $X \sim \mathrm{Poi}(\lambda)$이다.

(a) $\mathrm{Var}\, X \ge EX$이고, 등호는 $\mathrm{Var}\,\Lambda = 0$일 때뿐임을 보여라.

(b) $P(X=0) \ge e^{-\mu}$임을 보여라. 같은 평균의 Poisson보다 0이 많다.

(c) 노트 예제 6.3의 *may* 자료에서 0의 상대도수 $\frac{156}{262}$와 $e^{-0.656}$을 비교하고, (b)로 해석하라.

C8 음이항의 꼬리와 이항 ★

정리 1.3 · 15분

$Y \sim \mathrm{NB}(r, p)$이고 $B \sim \mathrm{b}(k+r, p)$이다.

(a) $P(Y \le k) = P(B \ge r)$임을 같은 사건을 두 가지로 읽어 보여라.

(b) $r = 3$, $p = \frac{1}{2}$, $k = 2$이면 $P(Y \le 2)$는 얼마인가?

C9 사건 수를 알 때의 위치 ★★

정의 3.1 · 20분

비율 λ의 Poisson 과정에서 $0 < s < t$이다.

(a) $P(T_1 \le s \mid N(t) = 1) = \frac{s}{t}$임을 보여라. 사건이 하나 있었다는 것을 알면 그 위치는 $(0, t)$ 위의 균등분포다.

(b) $N(t) = n$이 주어지면 $N(s) \sim \mathrm{b}(n, s/t)$임을 보여라. λ가 사라진다.

(c) $t = 1$, $s = 0.3$, $n = 5$이면 $P(N(s) = 2 \mid N(t) = 5)$는?

C10 $\Gamma(\frac{1}{2})$를 베타로 ★★

[T7] · 정리 4.2 · 20분

가우스 적분을 쓰지 않고 $\Gamma(\frac{1}{2}) = \sqrt{\pi}$를 증명한다.

(a) $x = \sin^2\theta$로 치환해 $B(\frac{1}{2}, \frac{1}{2}) = \int_0^1 x^{-1/2}(1-x)^{-1/2}dx$를 계산하라.

(b) 정리 4.2로 $\Gamma(\frac{1}{2})^2 = B(\frac{1}{2}, \frac{1}{2})\Gamma(1)$을 보이고 $\Gamma(\frac{1}{2}) = \sqrt{\pi}$를 얻어라. 양의 제곱근을 고르는 근거는?

(c) 거꾸로 $\int_{-\infty}^{\infty} e^{-z^2/2}dz = \sqrt{2\pi}$를 (b)에서 얻어라.

D

퀄형

여러 절과 앞 회차의 도구를 잇는 종합 문제다. 머리의 제한 시간을 재고, 노트를 덮고 푼다. 모형을 확률변수와 분포로 적는 것부터 답안의 일부다.

D1 두 Poisson 과정의 경주 ★★★

정리 2.4, 3.2, 3.3 · 1회 정리 1.10 · 제한 40분

종류 A의 사건이 비율 λ, 종류 B의 사건이 비율 μ인 독립인 Poisson 과정으로 일어난다.

(a) 두 과정을 합친 사건의 수가 비율 $\lambda+\mu$의 Poisson 과정임을 보여라.

(b) 첫 사건이 A일 확률이 $\dfrac{\lambda}{\lambda+\mu}$임을 보여라. 무기억성으로, 차례로 일어나는 사건들의 종류가 서로 독립이고 각각 확률 $\dfrac{\lambda}{\lambda+\mu}$로 A임을 설명하라.

(c) A가 n번 일어나는 것이 B가 m번 일어나는 것보다 먼저일 확률이 $P\big(\mathrm{b}(n+m-1, \dfrac{\lambda}{\lambda+\mu}) \geq n\big)$임을 보여라. 1회의 어떤 문제와 같은가?

(d) $\lambda=2$, $\mu=1$, $n=m=2$이면 이 확률은 얼마인가?

D2 사고를 잘 당하는 사람 ★★★

정리 4.5, 5.2, 5.3 · 3회 정리 2.5 · 제한 40분

Greenwood와 Yule(1920)의 모형이다. 한 노동자의 사고 위험률 Λ가 $\Gamma(r,\beta)$를 따르고, $\Lambda=\lambda$가 주어지면 첫해와 둘째 해의 사고 수 X_1, X_2는 독립인 $\mathrm{Poi}(\lambda)$이다.

(a) $\mathrm{Cov}(X_1,X_2)=r\beta^2$임을 보여라. X_1과 X_2는 독립인가?

(b) $X_1=k$가 주어지면 $\Lambda\sim\Gamma\!\left(r+k,\,\frac{\beta}{1+\beta}\right)$임을 보여라.

(c) $E(X_2\mid X_1=k)=\frac{(r+k)\beta}{1+\beta}$임을 보여라.

(d) $r=2,\ \beta=\frac{1}{2}$이면 $E(X_2\mid X_1=0)$은? (e) $E(X_2\mid X_1=3)$은?

D3 지수분포 표본의 평균 ★★★

정리 3.2, 3.7, 정의 3.8 · 제한 35분

$X_1, \dots, X_n$이 독립이고 평균 θ인 지수분포를 따른다. $\bar{X}$는 표본평균이다.

(a) $\sum_i X_i \sim \Gamma(n, \theta)$임을 보여라.

(b) $\frac{2n\bar{X}}{\theta} \sim \chi^2(2n)$임을 보여라. 이 분포는 θ에 의존하지 않는다.

(c) 정수 n과 $c > 0$에서 $P\big(\chi^2(2n) \le 2c\big) = P\big(\mathrm{Poi}(c) \ge n\big)$임을 보여라.

(d) $n = 5$이면 $P\big(\frac{2n\bar{X}}{\theta} \le 10\big)$은 얼마인가?

D4 기병대의 사후분포 ★★★

정리 4.5, 5.3 · 예제 2.6 · 제한 40분

위험률 Λ의 사전분포가 $\Gamma(\alpha, \beta)$이고, $\Lambda = \lambda$가 주어지면 $X_1, \dots, X_n$이 독립인 $\mathrm{Poi}(\lambda)$이다. $s = \sum_i x_i$로 쓴다.

(a) 사후분포가 $\Gamma\big(\alpha + s,\ \frac{\beta}{1+n\beta}\big)$임을 보여라.

(b) 사후평균이 사전평균 $\alpha\beta$와 표본평균 $\bar{x}$의 가중평균임을 보이고 가중치를 구하라.

(c) 다음 관측 X_{n+1}의 예측분포가 음이항분포임을 정리 5.3으로 보여라.

(d) 예제 2.6의 기병대 자료($n = 200$, $s = 122$)에 사전분포 $\Gamma(1, 1)$을 쓰면 사후평균은 얼마인가?

(e) 그때 다음 군단–해에 사망이 없을 예측확률은 얼마인가? $e^{-0.61}$과 비교하라.

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 6이 아닌 눈의 수는 $\mathrm{NB}(3, \frac{1}{6})$이다. 던지는 횟수는 그것에 3을 더한 것이다.

답 (a) $\binom{7}{5}\frac{5^5}{6^8} \approx 0.0391$ (b) 18

A2 힌트 $N=52, K=13, n=13$인 초기하분포다. 유한 모집단 수정 $\frac{N-n}{N-1}$을 잊지 않는다.

답 (a) $\frac{13}{4}$ (b) $\frac{507}{272} \approx 1.864$ (c) $\binom{39}{13}\big/\binom{52}{13} \approx 0.0128$

A3 힌트 (a) 다항계수 $\frac{10!}{3!\,3!\,3!\,1!}$. (c) 상관계수는 공분산을 두 표준편차로 나눈 것이고, 각 수의 주변분포는 이항이다.

답 (a) ≈ 0.0314 (b) -0.918 (c) ≈ -0.437

A4 힌트 $\lambda = np = 2$. (b)는 $1 - \sum_{k=0}^{2}\binom{1000}{k}(0.002)^k(0.998)^{1000-k}$이다.

답 (a) $1 - 5e^{-2} \approx 0.3233$ (b) ≈ 0.323323 (c) 0.004

A5 힌트 비율은 분당 $\lambda = 0.1$이다. 세 번째 전화의 시각은 $\Gamma(3, 1/\lambda)$이고, 첫 전화의 시각은 $\mathrm{Exp}(\lambda)$이다.

답 (a) $e^{-1} \approx 0.368$ (b) 30분 (c) 300분^2 (d) $e^{-1} - e^{-2} \approx 0.233$

A6 힌트 정리 2.5로 구매자 수와 비구매자 수는 독립이다. (c)에서는 전체가 고정되어 있으므로 이항이다.

답 (a) $e^{-3} \approx 0.0498$ (b) 3 (c) 2.1

A7 힌트 (a) $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$를 세 번. (b) $y = 2x$로 치환한다.

답 (a) $\frac{15\sqrt{\pi}}{8}$ (b) $\frac{3}{4}$ (c) $\frac{\sqrt{\pi}}{2}$

A8 힌트 (a)(b)는 $EX^k = \beta^k\Gamma(\alpha+k)/\Gamma(\alpha)$ (노트 연습 3.4). (c) X를 비율 $\frac{1}{2}$인 Poisson 과정의 세 번째 사건의 시각으로 읽는다.

답 (a) 48 (b) $\frac{1}{4}$ (c) $1 - \frac{5}{2}e^{-1} \approx 0.0803$

A9 힌트 $\chi^2(2) = \Gamma(1, 2)$는 평균 2인 지수분포다. (c)는 $\chi^2(3)$의 mgf를 $t = \frac{1}{4}$에서.

답 (a) $e^{-2} \approx 0.135$ (b) 6 (c) $2\sqrt{2}$

A10 힌트 밀도는 $12u^2(1-u)$이다. $B(3,2)=\dfrac{\Gamma(3)\Gamma(2)}{\Gamma(5)}=\dfrac{1}{12}$.

답 (a) $\frac{3}{5}$ (b) $\frac{1}{25}$ (c) $\frac{11}{16}$

A11 힌트 사후분포는 $\mathrm{Beta}(9,5)$다. (c)는 $E(p^2 \mid \text{자료})$이다. 두 시행은 p가 주어지면 독립이지만, p를 모르면 독립이 아니다.

답 (a) $\frac{9}{14}$ (b) $\frac{3}{196}$ (c) $\frac{3}{7}$

A12 힌트 (b) 집단 안의 분산의 평균과 집단 평균의 분산을 더한다. (c) 베이즈 정리.

답 (a) 3.1 (b) 4.99 (c) ≈ 0.896

A13 힌트 (a) $M \le \frac{1}{4}$은 다섯 점 가운데 3개 이상이 $\frac{1}{4}$ 이하라는 것과 같다. (b) $M \sim \mathrm{Beta}(3,3)$.

답 (a) $\frac{53}{512} \approx 0.104$ (b) $\frac{1}{28}$

A14 힌트 (b)에서 $P(X>x)=e^{-x/500}(1+\frac{x}{500})$이다. 이것은 비율 $\frac{1}{500}$인 Poisson 과정에서 x까지 사건이 1개 이하일 확률이다.

답 (a) $e^{-1}\approx 0.368$ (b) $2e^{-2}\approx 0.271$

B1 힌트 근거는 짧게 쓴다. 예: "정리 4.2(c)", "[T1] 지시함수". 공분산 항의 수는 순서 있는 쌍 (i,j), $i\neq j$의 수다.

B2 힌트 다섯째 단계에서 왜 "오른쪽"인지 생각한다. $G=1-F$이고 분포함수는 오른쪽 연속이다(1회 정리 6.3(c)).

B3 힌트 넷째 단계에서 $e^{-uv}e^{-(1-u)v}=e^{-v}$이고 $(uv)^{a-1}((1-u)v)^{b-1}\cdot v = u^{a-1}(1-u)^{b-1}v^{a+b-1}$이다.

B4 힌트 B1을 보지 않고 쓴다. 넷째 단계의 계산은 $\frac{K}{N}\left(\frac{K-1}{N-1}-\frac{K}{N}\right)$로 묶으면 짧다.

B5 힌트 유리수로 넓힐 때는 먼저 $\frac{1}{m}$에서의 값을 구한다. $G(1)=G(\frac{1}{m}+\cdots+\frac{1}{m})$.

B6 힌트 단계는 여섯이다. 역변환, 야코비안, 독립인 결합밀도, 변수변환 공식, 인수분해, 정규화 상수의 비교.

B7 힌트 지시변수의 분산은 $p(1-p)$이고, $i\neq j$인 순서쌍은 $n(n-1)$개다.

C1 힌트 상자 j가 비었다는 지시변수 $\mathbf{1}_j$로 $E=\sum_j \mathbf{1}_j$로 쓴다. 두 상자가 함께 비었을 확률은 모든 공이 나머지 $m-2$개 상자에 들어갈 확률이다.

답 (c) $10(0.9)^{10}\approx 3.487$ (d) ≈ 0.993

C2 힌트 (b) $\binom{k+r-1}{k} = \frac{r(r+1)\cdots(r+k-1)}{k!}$로 쓰고 $q_r^k = \frac{\lambda^k}{(r+\lambda)^k}$와 짝지운다. $p_r^r = (1+\lambda/r)^{-r}$이다.

C3 힌트 $P(N_1 = a_1, \dots, N_m = a_m) = P(N = \sum a_j)$· 다항 확률. $e^{-\lambda} = \prod_j e^{-\lambda p_j}$, $\lambda^{\sum a_j} = \prod_j \lambda^{a_j}$.

C4 힌트 (a) n번의 기회 각각이 "사건이 있고 표시됨", "사건이 있고 표시 안 됨", "사건 없음"의 세 범주로 독립적으로 나뉜다고 보면 된다. 그렇게 만들어도 같은 분포임을 결합 pmf로 확인한다.

답 (c) $-\frac{5}{8}$

C5 힌트 (b) $\{L > t, J = i\} = \{t < X_i < \min_{j\neq i} X_j\}$이다. $X_i = s$일 때 나머지가 모두 s보다 클 확률은 $e^{-(\Lambda-\lambda_i)s}$이다.

답 (c) $\frac{1}{3}$ (d) $\frac{1}{6}$

C6 힌트 (a) 오른쪽의 합에서 첨자를 $k \to k+1$로 옮긴다. (c) $g = \mathbf{1}_{\{k+1\}}$을 넣으면 $\lambda P(X=k) = (k+1)P(X=k+1)$이다. 그다음 k에 대한 귀납. `check()`에서는 λ를 `lam`으로 쓴다.

답 (b) $\lambda^3 + 3\lambda^2 + \lambda$

C7 힌트 (b) $P(X=0) = E\,e^{-\Lambda}$이고 $x \mapsto e^{-x}$는 볼록이다. Jensen 부등식 [T4].

C8 힌트 같은 시행열 위에서 "r번째 성공 전의 실패가 k개 이하"와 "처음 $k+r$번에 성공이 r개 이상"을 비교한다. 정리 3.2의 증명과 같은 생각이다.

답 (b) $\frac{1}{2}$

C9 힌트 $\{N(s)=k,\ N(t)=n\} = \{N(s)=k,\ N(t)-N(s)=n-k\}$이고 두 구간의 수는 독립이다(정의 3.1(i)).

답 (c) $\binom{5}{2}(0.3)^2(0.7)^3 = 0.3087$

C10 힌트 (a) $dx = 2\sin\theta\cos\theta\,d\theta$이고 $x^{-1/2}(1-x)^{-1/2} = \frac{1}{\sin\theta\cos\theta}$이다. (c)는 정리 3.5(c)의 증명을 거꾸로 읽는다.

답 (a) π

D1 힌트 (a) 정의 3.1의 두 조건을 확인한다. 같은 구간에서 두 과정의 수는 독립인 Poisson이다. (b) 첫 A와 첫 B의 시각은 독립인 지수분포다(3회 연습 3.4). (c) 처음 $n+m-1$개의 사건 가운데 A가 n개 이상인 것과 같다.

답 (d) $\frac{20}{27}$

D2 힌트 (a) $E[X_1X_2] = E[E(X_1X_2 \mid \Lambda)] = E\Lambda^2$. (b) 결합 "밀도–질량" $g(\lambda)P(X_1 = k \mid \lambda)$를 λ의 함수로 보고 감마 밀도의 꼴을 알아본다(정리 4.5와 같은 방법). (c) 반복기댓값.

답 (d) $\frac{2}{3}$ (e) $\frac{5}{3}$

D3 힌트 (b) $\Gamma(\alpha, \beta)$에 상수 $c > 0$을 곱하면 $\Gamma(\alpha, c\beta)$이다(mgf로 확인한다). (c) $\chi^2(2n) = \Gamma(n, 2)$를 비율 $\frac{1}{2}$인 Poisson 과정의 T_n으로 읽는다.

답 (d) $1 - \frac{523}{8}e^{-5} \approx 0.5595$

D4 힌트 (a) 정리 4.5와 같은 방법. $\lambda^{\alpha-1}e^{-\lambda/\beta}\prod_i e^{-\lambda}\lambda^{x_i}$. (c) 사후분포를 "혼합 분포"로 하는 Poisson의 혼합이다. NB(r, p)에서 $P(0) = p^r$.

답 (d) $\frac{123}{201} \approx 0.612$ (e) $(\frac{201}{202})^{123} \approx 0.5431$