

Probability and Statistics

PhD Qualification Exam

Mock Examination No. 2 – Solutions

Problem 1

10 points

除去零概率的对角线, (U, V) 的两个次序各贡献一次。令较大值为 m 、比值为 r , 则较小值为 mr , Jacobian 为 m 。因此

$$f_{M,R}(m, r) = 2m \mathbf{1}_{\{0 < m < 1, 0 < r < 1\}}.$$

积分得

$$f_M(m) = 2m \mathbf{1}_{(0,1)}(m), \quad f_R(r) = \mathbf{1}_{(0,1)}(r).$$

联合密度分解为 $f_M f_R$, 故 $M \perp R$, 且 $R \sim \text{Uniform}(0, 1)$ 。又 $\min\{U, V\} = MR$, 所以

$$\mathbb{E}(\min\{U, V\} \mid M) = M \mathbb{E}R = \frac{M}{2}.$$

Scoring guide (10 points). 变换、两块区域与 Jacobian 4 分, 支持 1 分; 两个边缘律 2 分、独立性 1 分; 条件期望 2 分。

Problem 2

10 points

对任意 $\varepsilon > 0$, 当 $n^\gamma > \varepsilon$ 时,

$$\mathbb{P}(|X_n| > \varepsilon) = \mathbb{P}(A_n) = \frac{1}{n} \rightarrow 0,$$

故 $X_n \xrightarrow{p} 0$ 。另一方面, $\sum_n \mathbb{P}(A_n) = \infty$ 且事件独立。第二 Borel–Cantelli lemma 给出 $\mathbb{P}(A_n \text{ i.o.}) = 1$ 。在这些指标上 $X_n = n^\gamma$, 所以 X_n 不可能几乎处处趋于零。

对 $p > 0$,

$$\mathbb{E}|X_n|^p = n^{\gamma p} \mathbb{P}(A_n) = n^{\gamma p - 1}.$$

因此 $X_n \rightarrow 0$ in L^p 当且仅当 $\gamma p < 1$ 。特别地

$$\mathbb{E}X_n = n^{\gamma - 1} \rightarrow \begin{cases} 0, & \gamma < 1, \\ 1, & \gamma = 1, \\ \infty, & \gamma > 1. \end{cases}$$

这也明确展示了 convergence in probability 并不蕴含 convergence of expectations。

Scoring guide (10 points). 依概率收敛 2 分; 第二 Borel–Cantelli 的条件与几乎处处结论 3 分; L^p 矩计算 3 分; 三种期望情形 2 分。

Problem 3**15 points**

取状态 0, 1, 2, 3 分别表示当前最长相关后缀为 empty、H、HH、HHT；状态 3 吸收。按状态顺序，转移矩阵为

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

这里从 HH 再得到 H 时，最长相关后缀仍是 HH。

令 $F_i(s)$ 为从状态 i 出发到吸收所需额外时间的 pgf, $F_3 = 1$ 。First-step analysis 给出

$$F_0 = \frac{s}{2}(F_0 + F_1), \quad F_1 = \frac{s}{2}(F_0 + F_2), \quad F_2 = \frac{s}{2}(F_2 + 1).$$

解得

$$F_2 = \frac{s}{2-s}, \quad \mathbb{E}[s^T] = F_0(s) = \frac{s^3}{(2-s)(4-2s-s^2)}, \quad 0 \leq s < 1.$$

该表达式连续延拓到 $s = 1$ 且值为 1。求导或对数求导，

$$\mathbb{E}T = F'_0(1) = 3 + \frac{1}{2-1} + \frac{2+2}{4-2-1} = 8.$$

Scoring guide (15 points). 状态定义 2 分、转移矩阵 3 分；三个 first-step 方程 3 分、求解 pgf 3 分；由 pgf 求均值 4 分。

Problem 4**15 points**

Poisson process 的独立平稳增量给出：在细分网格上，给定总到达数为 n 后，各到达落入网格单元的条件概率与单元长度成比例。令网格变细，得到 unordered 到达点为 n 个 i.i.d. $\text{Uniform}(0, T)$ 点；排序后

$$f(t_1, \dots, t_n \mid N_T = n) = \frac{n!}{T^n} \mathbf{1}_{\{0 < t_1 < \dots < t_n < T\}}.$$

若 $n \leq 1$ ，间隔条件自动成立。若 $n \geq 2$ 且 $0 \leq \delta \leq T/(n-1)$ ，作变换 $v_i = t_i - (i-1)\delta$ 。条件 $t_{i+1} - t_i \geq \delta$ 等价于

$$0 < v_1 < \dots < v_n < T - (n-1)\delta.$$

该单纯形体积为 $[T - (n-1)\delta]^n/n!$ ，故所求条件概率为

$$\left(1 - \frac{(n-1)\delta}{T}\right)^n.$$

若 $\delta > T/(n-1)$ ，概率为零。

令 C_δ 为距离不超过 δ 的 unordered pair 数。由 Poisson process 的二阶 Campbell formula（也可先给定 N_T 后计算）：

$$\mathbb{E}C_\delta = \lambda^2 \int_{0 < s < t < T} \mathbf{1}_{\{t-s \leq \delta\}} ds dt = \lambda^2 \left(T\delta - \frac{\delta^2}{2}\right).$$

Scoring guide (15 points). 条件 order-statistic 结论与证明 5 分；单纯形变换 3 分、体积 2 分、边界情形 1 分；pair-count 积分 2 分、结果 2 分。

Problem 5**15 points**

Brownian motion 几乎处处无界或由 reflection principle 可得 $\mathbb{P}(\tau_a > t) \rightarrow 0$, 故 $\tau_a < \infty$ a.s. 令 $k = \sqrt{2\lambda}$. Itô formula 表明

$$M_t = e^{-\lambda t} \cosh(kB_t)$$

是 martingale。对 $\tau_a \wedge n$ 使用 optional stopping; 停止前 $|B| \leq a$, 故停止后的变量由 $\cosh(ka)$ 控制, 可用 dominated convergence。于是

$$1 = \mathbb{E}M_{\tau_a} = \cosh(ka)\mathbb{E}e^{-\lambda\tau_a},$$

得到题设 Laplace transform。

在 $\lambda = 0$ 附近,

$$\operatorname{sech}(a\sqrt{2\lambda}) = 1 - a^2\lambda + \frac{5a^4}{6}\lambda^2 + o(\lambda^2).$$

右导数给出 moments; 其有限性也可由变换在零点邻域解析 (等价地, 退出有指数尾) 得到。因此

$$\mathbb{E}\tau_a = a^2, \quad \mathbb{E}\tau_a^2 = \frac{5a^4}{3}, \quad \operatorname{Var}(\tau_a) = \frac{2a^4}{3}.$$

路径映射 $B \mapsto -B$ 保持 Wiener measure, 不改变 τ_a , 却交换两个退出端点。因此对任意 Borel 集 $A \subset [0, \infty)$,

$$\mathbb{P}(\tau_a \in A, B_{\tau_a} = a) = \mathbb{P}(\tau_a \in A, B_{\tau_a} = -a) = \frac{1}{2}\mathbb{P}(\tau_a \in A).$$

所以退出方向与时间独立, 且 indicator 服从 Bernoulli(1/2)。

Scoring guide (15 points). 退出有限性 1 分、指数 martingale 2 分、optional stopping 条件与变换 2 分; 展开 2 分、两个矩与方差 3 分; 反射路径论证 3 分、独立性和 Bernoulli 律 2 分。

Problem 6**15 points**

Metropolis 规则给出

$$\begin{aligned} P(0,0) &= P(0,1) = \frac{1}{2}, \\ P(i,i+1) &= \frac{1}{2}, \quad P(i,i-1) = \frac{i}{2(i+1)}, \quad P(i,i) = \frac{1}{2(i+1)}, \quad 1 \leq i < m, \\ P(m,m-1) &= \frac{m}{2(m+1)}, \quad P(m,m) = \frac{m+2}{2(m+1)}. \end{aligned}$$

所有相邻转移均为正, 故链不可约; 每个状态有正 self-loop, 故非周期。对任意相邻 i, j , Metropolis 构造直接给出

$$\pi_i q(i,j) \min(1, \pi_j/\pi_i) = \pi_j q(j,i) \min(1, \pi_i/\pi_j).$$

因此 detailed balance 成立, π 平稳。有限不可约性保证平稳分布唯一。

Kac return-time formula 给出

$$\mathbb{E}_0\tau_0^+ = \frac{1}{\pi_0} = \frac{(m+1)(m+2)}{2}.$$

有限、不可约、非周期链的 ergodic theorem 给出

$$\lim_{n \rightarrow \infty} P^n(i, j) = \pi_j = \frac{2(j+1)}{(m+1)(m+2)}.$$

Scoring guide (15 points). 完整转移概率 5 分；不可约与非周期各 1 分、detailed balance 3 分、唯一平稳性 1 分；Kac 公式 2 分、极限 2 分。

Problem 7

10 points

令 $\bar{X} = S_n/n$ 。在 $0 < S_n < n$ 时，

$$\hat{\theta} = \frac{\bar{X}}{1 - \bar{X}}.$$

在闭包 $[0, \infty]$ 上， $S_n = 0$ 时 MLE 为 0， $S_n = n$ 时为 $+\infty$ 。

对固定内点 p ，CLT 给出 $\sqrt{n}(\bar{X} - p) \xrightarrow{d} N(0, p(1-p))$ 。取 $g(p) = p/(1-p)$ ， $g'(p) = (1-p)^{-2}$ ，Delta method 给出

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N\left(0, \frac{p}{(1-p)^3}\right) = N(0, \theta(1+\theta)^2).$$

对 $h(p) = \log[p/(1-p)]$ ，

$$\sqrt{n}\{\log \hat{\theta} - \log \theta\} \xrightarrow{d} N\left(0, \frac{1}{p(1-p)}\right).$$

令 $\widehat{\text{se}} = \{n\bar{X}(1-\bar{X})\}^{-1/2}$ 。先得到 $\log \hat{\theta} \pm z_{1-\alpha/2}\widehat{\text{se}}$ ，再指数化两个端点即得 odds 的渐近区间。

Scoring guide (10 points). MLE 及两个边界 3 分；CLT、Delta derivative 与方差 4 分；log-odds 方差、标准误和指数化区间 3 分。

Problem 8

10 points

在 H_0 下写共同速率为 λ 。Poisson splitting 给出 $X | S = s \sim \text{Binomial}(s, q_0)$ ，其中 $q_0 = t_1/(t_1 + t_2)$ ，与 λ 无关。

对每个 s ，选择整数 c_s 和 $\gamma_s \in [0, 1]$ 使

$$\mathbb{P}_{q_0}(X > c_s | S = s) + \gamma_s \mathbb{P}_{q_0}(X = c_s | S = s) = \alpha.$$

当 $X > c_s$ 时拒绝， $X = c_s$ 时以概率 γ_s 拒绝。特别地， $s = 0$ 时取 $\gamma_0 = \alpha$ 。于是对每个 s 的 conditional size 都是 α ，全期望公式保证对每个 nuisance rate 的 unconditional size 也是 α 。

在一般参数下， $X | S = s$ 是成功概率 $t_1\lambda_1/(t_1\lambda_1 + t_2\lambda_2)$ 的 $\text{Binomial}(s, \cdot)$ 变量。 $\lambda_1 > \lambda_2$ 时该成功概率严格大于 q_0 。Binomial family 在 X 中具有 monotone likelihood ratio，故 Karlin–Rubin theorem（或逐个 s 使用 Neyman–Pearson lemma）说明上述 upper-tail test 在每个条件样本空间内 UMP。因此它在所有 conditional-size- α tests 中 UMP。

Scoring guide (10 points). 条件二项分布与 nuisance 消除 3 分；随机化阈值 3 分、精确条件/无条件 size 1 分；备择下条件参数 1 分、MLR 与最优性 2 分。