信息学中的概率统计：作业二

截止日期：2025 年 10 月 10 日（周五）下课前。请务必通过教学网提交电子版。可下课前同时提交纸质版。

与本次作业相关的事宜，请发邮件给我(ruosongwang@pku.edu.cn)，抄送研究生助教叶昊洋(yhyfhgs@gmail.com)，以及负责本次作业的本科生助教周佳仪（flyfeather@stu.pku.edu.cn）。

第一题

(1) 对于任意 $a \geq 1$，构造**非负**离散随机变量 X，使得 $P(X \geq a \cdot E(X)) = 1/a$。

(2) 对于任意 $c \geq 1$，构造离散随机变量 X，使得 $P(|X - E(X)| \geq c \cdot \sigma(X)) = \frac{1}{c^2}$。

第二题

(1) 给定离散随机变量 X，证明对于任意实数 c，$E((X-c)^2) \geq \mathrm{Var}(X)$。

(2) 给定实数 a 和 b 满足 $a \leq b$，若离散随机变量 X 仅在 $[a, b]$ 上取值，也即 $P(a \leq X \leq b) = 1$，证明 $\mathrm{Var}(X) \leq \frac{(b-a)^2}{4}$。

第三题

在伯努利试验中，若 $P(A) = p$，令随机变量 X 为结果 A 首次出现时的试验次数，根据课上内容，有 $X \sim G(p)$。令随机变量 Y 为首次满足结果 A 和结果 $\overline{A}$ 均至少出现过一次时的试验次数。

(1) 写出 Y 的分布列，验证你的结果满足分布列的性质。

(2) 计算 $E(Y)$。

(3) 计算 $\mathrm{Var}(Y)$。不需要对结果进行化简。

第四题

令 $X \sim B(n, p)$。在课上，我们利用二项式系数的性质证明了 $E(X) = np$。在本题中，我们将用另一种方法计算 $E(X)$ 和 $E(X^2)$。

(1) 对于任意 $t \in \mathbb{R}$，计算 $E\left(e^{Xt}\right)$。

(2) 对于任意 $t \in \mathbb{R}$，证明

$$E\left(e^{Xt}\right) = \sum_{i=0}^{\infty} \frac{t^i}{i!} \cdot E(X^i)。$$

提示：对于固定的 $0 \leq k \leq n$，考虑对 e^{kt} 应用泰勒公式。

(3) 利用上一问中的结论，计算 $E(X)$ 和 $E(X^2)$。提示：令 $f(t) = E\left(e^{Xt}\right)$。如何利用上一问中的结论，通过 $f(t)$ 求得 $E(X)$ 和 $E(X^2)$?

(4) 令 $Y \sim \pi(\lambda)$，对于任意 $t \in \mathbb{R}$，计算 $E\left(e^{Yt}\right)$。在此基础上，参考上一问中的方法，计算 $E(Y)$ 和 $E(Y^2)$。

第五题

在课上，我们考虑了如下球与桶模型：有 n 个球，每个球都等可能被放到 m 个桶中的任一个。在本题中，我们考虑 $m = n$ 的情况，并假设 $n = m \geq 2$。在本题中，log 表示以 2 为底的对数。

(1) 随机变量 X_i 表示第 i 个桶中球的数量。对于任意 $i \in \{1, 2, \ldots, n\}$，证明 $E(X_i) = 1$。

(2) 对于任意 $i \in \{1, 2, \ldots, n\}$ 和任意 $1 \leq k \leq n$，证明 $P(X_i = k) \leq \frac{1}{k!}$。

(3) 定义随机变量 $Y = \max\{X_1, X_2, \ldots, X_n\}$，证明 $P(Y \geq 4 \log n) \leq 1/n$。提示：考虑使用 Union Bound。

(4) 证明 $E(Y) \leq 5 \log n$。

第六题

给定离散随机变量 X，假设其期望 $E(X)$ 和标准差 $\sigma(X)$ 均存在。

(1) 对于任意实数 b，证明 $E((X - E(X) + b)^2) = (\sigma(X))^2 + b^2$。

(2) 对于任意实数 $t > \sigma(X)$，证明 $P(X \geq E(X) + t) < 1/2$。提示：参考切比雪夫不等式的证明方法，利用上一问中的结论并选取合适的 b。

(3) 对于任意实数 m，若满足 $P(X \geq m) \geq 1/2$ 且 $P(X \leq m) \geq 1/2$，利用上一问中的结论，证明 $|E(X) - m| \leq \sigma(X)$。