信息学中的概率统计：作业三

截止日期：2025 年 10 月 24 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卢宇宸（lyc1091864029@stu.pku.edu.cn）。

第一题

(1) X 为离散随机变量，且 X 仅取非负整数值。证明 $E(X) = \sum_{x=0}^{+\infty} P(X > x)$。

(2) X 为连续随机变量，且 X 仅取非负实数值。证明 $E(X) = \int_0^{+\infty} P(X > x)dx$。

第二题

连续随机变量 X 的分布函数 $F(x)$ 为严格单调增的连续函数，其反函数存在。

(1) 令 $Z \sim U(0,1)$。证明 $Y = F^{-1}(Z)$ 服从与 X 相同的分布，且 $W = F(X)$ 服从与 Z 相同的分布。

(2) 利用上一问中的结论构造函数 f，使得 $f(Z)$ 服从参数为 λ 的指数分布。这里 $Z \sim U(0,1)$。

第三题

给定 $[0,1]$ 上的 n 个区间 $[l_i, r_i]$，满足 $0 \le l_i < r_i \le 1$。

(1) 令 $X \sim U(0,1)$，定义函数 $g(x) = |\{i \mid x \in [l_i, r_i]\}|$，也即包含 x 的区间数量。计算 $E(g(X))$。

(2) 证明存在 $x \in [0,1]$，它被至少 $m = \lceil \sum_{i=1}^{n} (r_i - l_i) \rceil$ 个区间所包含。

第四题

对于实数参数 μ 和 $b > 0$，已知连续随机变量 X 的概率密度函数满足对于任意 $x \in \mathbb{R}$，

$$f(x) = c \cdot e^{-|x-\mu|/b},$$

这里 c 为与参数 μ 和 b 有关的常数。

(1) 计算常数 c 以及 X 的分布函数

(2) 计算 $E(X)$ 和 $\mathrm{Var}(X)$

第五题

(1) 对于任意实数 $x>0$，证明

$$\int_x^{+\infty} \frac{t}{x} e^{-t^2/2} dt = \frac{e^{-x^2/2}}{x}。$$

(2) 令

$$g(x) = \left(\int_x^{+\infty} e^{-t^2/2} dt\right) - \frac{x}{x^2+1} e^{-x^2/2}。$$

证明当 $x \geq 0$，$g(x) \geq 0$。

(3) 令 $X \sim N(0,1)$，证明对于任意实数 $x>0$,

$$\frac{x}{x^2+1} \cdot \frac{e^{-x^2/2}}{\sqrt{2\pi}} \leq P(X \geq x) \leq \frac{1}{x} \cdot \frac{e^{-x^2/2}}{\sqrt{2\pi}}。$$

(4) 令 $Y \sim N(\mu, \sigma^2)$，证明对于任意实数 $k>0$,

$$1 - \frac{k}{k^2+1} \cdot e^{-k^2/2} \cdot \sqrt{\frac{2}{\pi}} \geq P(|Y-\mu| \leq k\sigma) \geq 1 - \frac{1}{k} \cdot e^{-k^2/2} \cdot \sqrt{\frac{2}{\pi}}。$$

第六题

(1) 令 $X \sim \Gamma(\alpha, \lambda)$。对于任意实数 t，计算 $E(e^{tX})$。

(2) 若 $Y \sim N(0,1)$，对于任意实数 t，计算 $E(e^{tY^2})$。

(3) 若 $Y \sim N(0,1)$，计算 $Z = Y^2$ 的概率密度函数，并将结果与伽玛分布和 χ^2 分布的概率密度函数相比较。