信息学中的概率统计：作业三

截止日期：2024 年 11 月 1 日（周五）下课前。**如无特殊情况，请不要提交电子版！**

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 $E(X) = \int_0^{+\infty} P(X > x)dx$。

答案：

(1)

$$
\begin{aligned}
&\sum_{x=0}^{+\infty} P(X > x) \\
=&\sum_{x=0}^{+\infty} \sum_{y=x+1}^{\infty} P(X = y) \\
=&\sum_{y=1}^{\infty} P(X = y) \cdot |\{x \mid y \geq x+1, x \geq 0\}| \\
=&\sum_{y=1}^{\infty} P(X = y) \cdot y \\
=&\sum_{y=0}^{\infty} P(X = y) \cdot y = E(X)
\end{aligned}
$$

(2)

$$
\begin{aligned}
&\int_0^{+\infty} P(X > x)dx \\
=&\int_0^{+\infty} \int_x^{+\infty} f(y)dydx \\
=&\int_0^{\infty} f(y) \int_0^{y} dxdy \\
=&\int_0^{\infty} f(y)ydy \\
=&E(X)
\end{aligned}
$$

第二题

在 Unix 操作系统中，用随机变量 X 表示一个随机的任务所需的内存。历史数据表明，对于任意实数 $x \geq 1$，$P(X > x) = 1/x^{\alpha}$。这里 $\alpha \in (0, 2)$ 是固定常数。

(1) 计算随机变量 X 的概率分布函数和概率密度函数。

(2) 计算 $E(X)$ 和 $E(X^2)$

答案:

(1) $F(X) = P(X \leq x) = 1 - P(X > x) = 1 - x^{-\alpha}$。$f(x) = F'(x) = \alpha x^{-\alpha-1}$。

(2) 注意到

$$E(X) = \int_1^\infty \alpha x^{-\alpha-1} x dx = \int_1^\infty \alpha x^{-\alpha} dx = \begin{cases} \frac{\alpha}{\alpha-1} & \alpha > 1 \\ +\infty & \alpha \in (0,1] \end{cases},$$

另外

$$E(X^2) = \int_1^\infty f(x) x^2 dx = \int_1^\infty \alpha x^{-\alpha+1} dx = +\infty。$$

第三题

(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 $X \sim N(0,1)$，也即连续随机变量 X 服从标准高斯分布，证明对于任意实数 $x > 0$,

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

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

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

答案:

(1) 注意到

$$\int t e^{-t^2/2} dt = -e^{-t^2/2} + C,$$

因此

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

也即

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

(2)

$$P(X \geq x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-t^2/2} dt \leq \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} \frac{t}{x} e^{-t^2/2} dt = \frac{e^{-x^2/2}}{x\sqrt{2\pi}}。$$

(3) 令 $Z = (Y - \mu)/\sigma$，则 $Z \sim N(0,1)$。根据第二问的结论，对于任意实数 k,

$$P(Z \geq k) \leq \frac{e^{-k^2/2}}{k\sqrt{2\pi}}$$

也即

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

另外，

$$P(Z \leq -k) \leq \frac{e^{-k^2/2}}{k\sqrt{2\pi}}$$

也即

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

因此，

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

也即

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

第四题

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。证明 $Y = F(X)$ 服从 $(0,1)$ 上的均匀分布 $U(0,1)$。

答案： 因为 $0 \leq F(x) \leq 1$，因此 Y 仅在 $[0,1]$ 上取值。若 $F_Y(y)$ 表示 Y 的分布函数，则当 $y < 0$ 时，$F_Y(y) = 0$。当 $y \geq 1$ 时，$F_Y(y) = 1$。当 $0 \leq y < 1$，

$$F_Y(y) = P(Y \leq y) = P(F(X) \leq y) = P(X \leq F^{-1}(y)) = F(F^{-1}(y)) = y。$$

因此

$$F_Y(y) = \begin{cases} 0 & y < 0 \\ y & 0 \leq y < 1 \\ 1 & y \geq 1 \end{cases},$$

也即 $Y \sim U(0,1)$。

第五题

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 $\mathrm{Var}(X)$

答案：

(1)

$$\int_{-\infty}^{+\infty} e^{-|t-\mu|/b} dt = \int_{-\infty}^{+\infty} e^{-|t|/b} dt = 2\int_{0}^{+\infty} e^{-t/b} dt = 2b$$

因此 $c = \frac{1}{2b}$。

对于 $x \le \mu$,

$$P(X \le x) = \frac{1}{2b}\int_{-\infty}^{x} e^{-|t-\mu|/b}dt = \frac{1}{2b}\int_{-\infty}^{x} e^{(t-\mu)/b}dt = \frac{1}{2}e^{(x-\mu)/b}$$

对于 $x \ge \mu$,

$$P(X \le x) = 1 - P(X \ge x) = 1 - \frac{1}{2b}\int_{x}^{+\infty} e^{-|t-\mu|/b}dt = 1 - \frac{1}{2b}\int_{x}^{+\infty} e^{-(t-\mu)/b}dt = 1 - \frac{1}{2}e^{-(x-\mu)/b}$$

(2)

$$E(X) = \frac{1}{2b}\int_{-\infty}^{+\infty} t \cdot e^{-|t-\mu|/b}dt = \frac{1}{2b}\ \int_{-\infty}^{+\infty} (t+\mu) \cdot e^{-|t|/b}dt = \mu$$

$$E(X^2) = \frac{1}{2b}\int_{-\infty}^{+\infty} t^2 \cdot e^{-|t-\mu|/b}dt = \frac{1}{2b}\int_{-\infty}^{+\infty} (t+\mu)^2 \cdot e^{-|t|/b}dt = \frac{1}{2b}\int_{-\infty}^{+\infty} (t^2+\mu^2) \cdot e^{-|t|/b}dt$$

$$\frac{1}{2b}\int_{-\infty}^{+\infty} t^2 \cdot e^{-|t|/b}dt = \frac{1}{b}\int_{0}^{+\infty} t^2 \cdot e^{-t/b}dt = b^2\int_{0}^{+\infty} t^2 \cdot e^{-t}dt = \Gamma(3) \cdot b^2 = 2b^2$$

因此,

$$E(X^2) = 2b^2 + \mu^2,$$

也即 $\mathrm{Var}(X) = 2b^2$

第六题

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

(2) 对于正整数 n，若 $Y_n \sim \chi^2(n)$，也即 $Y_n \sim \Gamma(n/2, 1/2)$。对于任意实数 t，计算 $E(e^{tY_n})$

(3) 若 $X \sim N(0,1)$，计算 $Y = X^2$ 的概率密度函数

答案:

(1)

$$E(e^{tX^2}) = \int_{-\infty}^{+\infty} e^{tx^2} \cdot \frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-(1/2-t)x^2}dx$$

当 $t \ge 1/2$ 时，积分发散，因此期望不存在。当 $t < 1/2$ 时，令 $y = \sqrt{1-2t}x$，则有 $-\frac{y^2}{2} = -x^2(1/2-t)$，因此

$$\int_{-\infty}^{+\infty} e^{-(1/2-t)x^2}dx = \frac{1}{\sqrt{1-2t}}\int_{-\infty}^{+\infty} e^{-y^2/2}dy$$

也即

$$E(e^{tX^2}) = \frac{1}{\sqrt{1-2t}}。$$

(2)

$$E(e^{tY_n}) = \frac{1}{2^{n/2}} \cdot \frac{1}{\Gamma(n/2)} \cdot \int_{0}^{+\infty} y^{n/2-1} \cdot e^{-(1/2-t)y}dy。$$

当 $t \ge 1/2$ 时，积分发散，因此期望不存在。当 $t < 1/2$ 时, 令 $z = (1/2-t)y$,

$$\begin{aligned}&\frac{1}{2^{n/2}} \cdot \frac{1}{\Gamma(n/2)} \cdot \int_{0}^{+\infty} y^{n/2-1} \cdot e^{-(1/2-t)y}dy \\ =&\frac{1}{2^{n/2}} \cdot \frac{1}{\Gamma(n/2)} \cdot (1/2-t)^{-n/2}\int_{0}^{+\infty} z^{n/2-1} \cdot e^{-z}dy = (1-2t)^{-n/2}。\end{aligned}$$

(3) 令 $Y = X^2$，则有 Y 仅取非负实数。对于 $y \geq 0$

$$P(Y \leq y) = P(X^2 \leq y) = P(X \in [-\sqrt{y}, \sqrt{y}]) = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} \cdot e^{-t^2/2} dt = 2\int_{0}^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} \cdot e^{-t^2/2} dt$$

用 f_Y 表示 Y 的概率密度函数，则有对于任意 $y \geq 0$，

$$f_Y(y) = 2 \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-y/2} \cdot \frac{1}{2\sqrt{y}} = \frac{1}{\sqrt{2\pi y}} e^{-y/2},$$

也即 $Y \sim \chi^2(1)$