实验物理中的统计方法 作业2

1.

习题 1.6. 假设随机变量 x 的概率密度函数为 $f(x)$。证明 $y=x^2$ 的概率密度函数为

$$g(y)=\frac{1}{2\sqrt{y}}f(\sqrt{y})+\frac{1}{2\sqrt{y}}f(-\sqrt{y}). \tag{1.2}$$

证明：

由概率守恒：

$$g(y)\mathrm{d}y=f(x)\mathrm{d}x$$

$$g(y)=f(\sqrt{y})\cdot\left|\frac{\mathrm{d}\sqrt{y}}{\mathrm{d}y}\right|+f(-\sqrt{y})\cdot\left|\frac{\mathrm{d}-\sqrt{y}}{\mathrm{d}y}\right|$$

化简得到

$$g(y)=\frac{1}{2\sqrt{y}}f(\sqrt{y})+\frac{1}{2\sqrt{y}}f(-\sqrt{y})$$

证毕。

2.

习题 1.7. 假设两个独立的随机变量 x 和 y 都服从 *0* 到 *1* 之间的均匀分布，即概率密度函数 $g(x)$ 为

$$g(x)=\begin{cases}1 & 0<x<1\\ 0 & \text{其它},\end{cases} \tag{1.3}$$

概率密度函数 $h(y)$ 与 $g(x)$ 类似。

(a) 利用 *Statistical Data Analysis* 中的 *(1.35)* 式，证明，$z=xy$ 的概率密度函数 $f(z)$ 为

$$f(z)=\begin{cases}-\log z & 0<z<1,\\ 0 & \text{其它}.\end{cases} \tag{1.4}$$

(b) 利用 *Statistical Data Analysis* 的 *(1.37)* 和 *(1.38)* 式，通过另外定义一个函数 $u=x$，求 $z=xy$ 的概率密度函数。首先求 z 和 u 的联合概率密度函数，然后对 u 进行积分求出 z 的概率密度函数。

(c) 证明 z 的累积分布为

$$F(z)=z(1-\log z). \tag{1.5}$$

解：

(a)

$$f(z)=\int_{-\infty}^{\infty}g(\frac{z}{y})h(y)\frac{\mathrm{d}y}{|y|}$$

被积函数不为零要求：

$$z<y<1,\ z>0$$

故只有$0<z<1$时由不为零的值。

即：

$$f(z)=\int_z^1\frac{1}{y}\mathrm{d}y=-\log z$$

(b)

据题有:

$$g(u,z)=f(x,y)|J|$$

其中

$$x=u, y=\frac{z}{u}$$

$$J=\det\begin{pmatrix}\frac{\partial x}{\partial u} & \frac{\partial x}{\partial z}\\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial z}\end{pmatrix}$$

得到:

$$J=\det\begin{pmatrix}1\,0\\ \frac{-z}{u^2} & \frac{1}{u}\end{pmatrix}=\frac{1}{u}$$

有:

$$f(u,z)=g(u)h(z/u)\cdot\frac{1}{|u|}$$

积分得到:

$$f(z)=\begin{cases}-\log z, 0<z<1\\ 0, \text{else.}\end{cases}$$

(c)

$$F(z)=\int_0^z -\log z'\mathrm{d}z'=z(1-\log z)$$

3.

习题 1.8. 考虑随机变量 x 与常数 α 和 β。证明

$$\begin{aligned}E[\alpha x+\beta]&=\alpha E[x]+\beta,\\ V[\alpha x+\beta]&=\alpha^2 V[x].\end{aligned} \tag{1.6}$$

证明:

（1）设概率密度函数为$f(x)$

$$E[\alpha x+\beta]=\int(\alpha x+\beta)f(x)\mathrm{d}x=\alpha\int xf(x)\mathrm{d}x+\beta\int f(x)\mathrm{d}x=\alpha E[x]+\beta$$

（2）

$$\begin{aligned}V[\alpha x+\beta]&=\int(\alpha x+\beta-E[\alpha x+\beta])^2 f(x)\,\mathrm{d}x\\ &=\int(\alpha x+\beta)^2 f(x)\,\mathrm{d}x-E[\alpha x+\beta]\\ &=\alpha^2E[x^2]+2\alpha\beta E[x]\mathrm{d}x+\beta^2-(\alpha E[x]+\beta)^2\\ &=\alpha^2(E[x^2]-E[x]^2)\\ &=\alpha^2V[x]\end{aligned}$$

4.

习题 1.10. 假设随机变量 $\mathbf{x} = (x_1, \ldots, x_n)^T$ 用联合概率密度函数$f(\mathbf{x})$ 描述，而变量 $\mathbf{y} = (y_1, \ldots, y_n)^T$ 由下面的线性变换定义

$$y_i = \sum_{j=1}^{n} A_{ij} x_j. \tag{1.8}$$

假设反变换 $\mathbf{x} = A^{-1}\mathbf{y}$ 存在。

(a) 证明 $\mathbf{y}$ 的联合概率密度函数为

$$g(\mathbf{y}) = f(A^{-1}\mathbf{y})|\det(A^{-1})|. \tag{1.9}$$

(b) 当 A 为矩阵，即 $A^{-1} = A^T$ 时，求 $g(\mathbf{y})$。

(a)

$$g(y) = f(A^{-1}y)|J|$$

其中：

$$J = \frac{D(x_1, \ldots, x_n)}{D(y_1, \ldots, y_n)}$$

设矩阵A^{-1}的矩阵元为a_{ij}，则

$$\frac{\partial x_i}{\partial y_j} = a_{ij}$$

故

$$J = \det(A^{-1})$$

得到

$$g(y) = f(A^{-1}y)|\det(A^{-1})|$$

(b)

由题，A为正交矩阵，故行列式为± 1

$$g(y) = f(A^{-1}y)$$

5.

习题 1.12. A 和 B 两人分别轮流掷一枚硬币，最先掷得正面朝上的人胜出，假定 A 先开始掷。

(a) 如果所掷的硬币是公平硬币，A 胜出的概率是多少？

(b) 假设 $P(\text{正面朝上}) = p$，p 可能不等于 $\frac{1}{2}$。A 胜出的概率是多少？

(c) 证明：对任意的 p，$0 < p < 1$，$P(A\text{ 胜出}) > \frac{1}{2}$。

(a)

$$P(A) = \frac{1}{2} + (\frac{1}{2})^3 + (\frac{1}{2})^5 + \ldots = \frac{1}{2} \times \frac{1}{1 - \frac{1}{4}} = \frac{2}{3}$$

(b)

$$P(A) = p + p(1-p)^2 + p(1-p)^4 + \ldots = p \times \frac{1}{1 - (1-p)^2} = \frac{1}{2-p}$$

(c)

由$0 < p < 1$已知，$P(A) > \frac{1}{2}$。

6.

习题 1.13. 假设 X 和 Y 是两个连续的随机变量，且其方差有限。证明相关系数 $\rho \equiv \frac{\mathrm{cov}(X,X)}{\sigma_X\sigma_Y} = \pm 1$ 的充要条件是 X 与 Y 几乎处处有线性关系，即存在常数 $a \neq 0$ 和 b 使得 $Y = aX + b$。

先证充分性：

$$\rho = \frac{E[XY] - E[X]E[Y]}{\sqrt{E[X^2] - E[X]^2}\sqrt{E[Y^2] - E[Y]^2}}$$

若

$$Y = aX + b$$

则

$$\rho = \frac{aE[X^2] - aE[X]^2}{|aE[X^2] - aE[X]^2|} = \mathrm{sgn}a$$

再证必要性：

设μ_x, μ_y为X, Y的平均值，则由定义：

$$\mathrm{cov}(X,Y) = \int (x-\mu_x)(y-\mu_y)f(x,y)\mathrm{d}x\mathrm{d}y$$

$$\sigma_X^2 = \int (x-\mu_x)^2 f(x,y)\mathrm{d}x\mathrm{d}y$$

$$\sigma_Y^2 = \int (y-\mu_y)^2 f(x,y)\mathrm{d}x\mathrm{d}y$$

由柯西-施瓦兹不等式可知，

$$cov(X,Y)^2 \leq \sigma_X^2\sigma_Y^2$$

当且仅当$(x-\mu_x) = k(y-\mu_y)$几乎处处成立时等号成立，即Y与X成线性关系。

> 柯西-施瓦兹不等式证明：
>
> $$\int ((x-\mu_x) + a(y-\mu_y))^2 f(x,y)\mathrm{d}x\mathrm{d}y = \sigma_x^2 + 2a\mathrm{cov}(X,Y) + \sigma_Y^2 \geq 0$$
>
> 由a任意知：
>
> $$\mathrm{cov}(X,Y)^2 \leq \sigma_X^2\sigma_Y^2$$
>
> 由二次函数的性质知，等号成立说明上面的积分式可以取值为0，故知$\exists a = k$，使得$(x-\mu_x) = k(y-\mu_y)$几乎处处成立。证毕。

7.

习题 1.14. 设 (X,Y) 是二维连续随机变量，且，且 $E(X)$ 存在。证明：$E(X) = E[E[X|Y]]$，其中 $E[X|Y]$ 是给定 Y 的条件下 X 的数学期望。

由定义：

$$
\begin{aligned}
E[E[X|Y]] &= \iint E[X|Y]f(x,y)\mathrm{d}x\mathrm{d}y \\
&= \iint \left(\frac{\int xf(x,y)\mathrm{d}x}{\int f(x,y)\mathrm{d}x}\right) f(x,y)\mathrm{d}x\mathrm{d}y \\
&= \int \left(\frac{\int xf(x,y)\mathrm{d}x}{\int f(x,y)\mathrm{d}x}\right)(\int f(x,y)\mathrm{d}x)\mathrm{d}y \\
&= \int xf(x)\mathrm{d}x\mathrm{d}y \\
&= E[X]
\end{aligned}
$$

证毕。

8.

习题 1.15. 考虑两个连续随机变量 X 和 Y，其联合概率密度函数为

$$
f(x,y) = \begin{cases} \frac{1}{\pi R^2}, & x^2+y^2 \le R^2, \\ 0, & \text{其他} \end{cases}
$$

(a) 求边缘概率密度函数 $f_X(x)$ 和 $f_Y(y)$。

(b) 求条件概率密度函数 $f(x|Y=y)$ 和 $f(y|X=x)$。

(c) 验证所得到的条件概率密度函数和边缘概率密度函数满足贝叶斯定理。

(d) 判断 X 和 Y 是否相互独立。

解:

(a)

$$
f_X(x) = \int f(x,y)\mathrm{d}y = \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \frac{1}{\pi R^2}\mathrm{d}y = \frac{2\sqrt{R^2-x^2}}{\pi R^2}
$$

由X与Y的对称性，得到:

$$
f_Y(y) = \frac{2\sqrt{R^2-y^2}}{\pi R^2}
$$

(b)

$$
f(x|Y=y) = \frac{f(x,y)}{\int f(x,y)\mathrm{d}x} = \frac{1}{2\sqrt{R^2-y^2}}
$$

由对称性:

$$
f(y|X=x) = \frac{1}{2\sqrt{R^2-x^2}}
$$

(c)

$$
f(x|Y=y) = \frac{1}{2\sqrt{R^2-y^2}} = \frac{f(x,y)}{f_Y(y)}
$$

$$
f(y|X=x) = \frac{1}{2\sqrt{R^2-x^2}} = \frac{f(x,y)}{f_X(x)}
$$

贝叶斯定理成立。

(d)

$$f_X(x)f_Y(y) = \frac{2\sqrt{R^2 - x^2}}{\pi R^2} \cdot \frac{2\sqrt{R^2 - y^2}}{\pi R^2}$$

而

$$f(x, y) = \frac{1}{\pi R^2}$$

故

$$f(x, y) \neq f_X(x)f_Y(y)$$

所以X与Y不相互独立。