作业 **1**	球均匀分布

本次作业基本是初等概率计算，其中 3-4 题为第二讲多元球对称分布的简单情况（预热，两周后提交）。

背景：几何概率问题中经常会碰到“在一个集合中随机选取一个点”这样的表述，这里的“随机 (at random)”指的是处处等可能，即均匀的意思。下面我们介绍一般的均匀分布。n 维随机向量 $\mathbf{x}$ 服从集合 $S \subset R^n$ 上的均匀分布，记作 $\mathbf{x} \sim U(S)$, 如果对任何 $S_0 \subset S$,

$$P(\mathbf{x} \in S_0) = |S_0|/|S|,$$

其中 $|S_0|$ 表示集合 S_0 的大小 (size, 个数、长度、面积、体积等). 例如 $B^2 = \{(x_1, x_2) : x_1^2 + x_2^2 \le 1\}$ 为平面上的单位圆，随机向量 $\mathbf{x} = (x_1, x_2)^\top$ 服从 $\sim U(B^2)$ 当且仅当

$$P(\mathbf{x} \in S_0) = |S_0|/\pi, S_0 \subset B^2$$

其概率密度为

$$f(x_1, x_2) = \begin{cases} \frac{1}{\pi}, & x_1^2 + x_2^2 \le 1 \\ 0, & x_1^2 + x_2^2 > 1. \end{cases}$$

再如，$\mathbf{x} = (x_1, x_2)^\top$ 服从平面单位圆周 $S^1 \subset R^2$ 上的均匀分布, 记作 $\mathbf{x} \sim U(S^1)$ 当且仅当对任何圆弧 $L \subset S^1$

$$P(\mathbf{x} \in L) = |L|/(2\pi)$$

其中 $|L|$ 为弧长。该分布关于 R^2 的 Lebesgue 测度不存在概率密度。

1. R^1 中原点为中心的单位球为区间 $B^1 = [-1, 1]$, 球面由两个点构成：$S^0 = \{-1, 1\}$。
   (a) 假设随机变量 x 服从球内均匀分布，即 $x \sim U(-1, 1)$, 证明半径 $r = |x| \sim U(0, 1)$
   (b) 令 $u = x/|x|$, 证明 $u \sim U(S^0)$ (球面均匀分布), 即 $P(u = \pm 1) = 1/2$。
2. (a) 假设二元随机向量 $(x, y) \sim U(B^2)$, 即有概率密度

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

   试求 x 的边际分布。
   (b) 假设二元随机向量 $(x, y) \sim U(S^1)$ (原点为中心的单位圆周上的均匀分布)，试求 x 的边际分布。
3. (a) 假设 x 是关于 0 对称的随机变量，$P(x = 0) = 0$, 令 $u = x/|x|$, 证明 $P(u = \pm 1) = 1/2$, $u \perp\!\!\!\perp |x|$.
   (b) 条件同 (a), 假设 $\sigma \perp\!\!\!\perp x$, $\sigma \sim U(S^0)$ 即 $P(\sigma = \pm 1) = 1/2$, 证明 $x \overset{d}{=} x\sigma \overset{d}{=} |x|\sigma$。
   (c) 假设 x, y 都是关于 0 对称的独立随机变量，$P(y = 0) = 0$, 证明 $x/y \overset{d}{=} x/|y|$。
   (d) 假设 x, y iid $\sim N(0, 1)$, 证明 $x/y \sim t_1$(自由度为 1 的 t 分布, 也是 Cauchy 分布)。

4. 假设二元随机向量 $\mathbf{x} = (x_1, x_2)^\top \in R^2$ 的概率密度函数 $f(\mathbf{x})$ 仅与 $\|\mathbf{x}\|$ 有关，即存在某个一元函数 h 使得

$$f(\mathbf{x}) = h(\|\mathbf{x}\|)$$

这意味着在正交旋转下 $\mathbf{x}$ 的分布不变，此时我们称 $\mathbf{x}$ 服从球对称分布。令极坐标变换

$$x_1 = r\cos(\theta), x_2 = r\sin(\theta), \quad r > 0,\ 0 \le \theta < 2\pi,$$

记 r 的概率密度函数为 $p(r)$。

(a) 试证明 r 与 θ 独立，且 $\theta \sim U(0, 2\pi)$, r 的概率密度 $p(r) = 2\pi r h(r)$。

(b) 若 $p(r) = 2r1_{(0<r<1)}$, 证明 $\mathbf{x} \sim U(B^2)$。

(c)（Box-Muller transform）假设 r 服从自由度为 2 的 χ 分布（记作 $r \sim \chi_2$，也称为 Rayleigh 分布）, 即 r 的概率密度函数为

$$p(r) = re^{-r^2/2}, r > 0,$$

证明 x_1, x_2 iid $\sim N(0,1)$。(注：Box-Muller transform 方法是标准正态随机数最常用的产生方法，其中 $r \sim \chi_2$ 可由均匀分布随机数 $U \sim U(0,1)$ 生成如下 (inversion sampling)：$r = \sqrt{-2\log(U)}$.)

(d) 若 r 服从 $U(0,1)$, 即 $p(r) = 1, 0 < r < 1$，求 $f(\mathbf{x})$，它在球心处的密度是多少？

5. n 元函数 $f(\mathbf{x}), \mathbf{x} \in R^n$ 是正交不变的（即对任何正交矩阵 H, $f(H\mathbf{x}) = f(\mathbf{x})$) 当且仅当 f 具有形式 $f(\mathbf{x}) = h(\|\mathbf{x}\|)$, 其中 $h(\cdot)$ 是一元函数。

6. 单位球内均匀分布 $U(B^n)$ 的 $k < n$ 维边际是否可能服从均匀分布 $U(B^k)$? 简单说明理由。