# Homework 1 参考答案

**Exercise 1.1.** 假设二元随机向量 $(x, y) \sim U(B_2)$，即有概率密度

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

(a) 证明 $x$ 服从半圆律（即 $x$ 的概率密度为 $p(t) = \frac{2}{\pi}\sqrt{1-t^2}$，$|t| \le 1$）。

(b) 证明 $(x+1)/2 \sim \mathrm{Beta}(3/2,\, 3/2)$。

(c) 证明 $x^2 \sim \mathrm{Beta}(1/2,\, 3/2)$。

(d) 证明 $\sqrt{3}\, x/\sqrt{1-x^2} \sim t_3$。

（提示：利用性质：若 $x \sim \chi^2_n$，$y \sim \chi^2_m$，$x \perp\!\!\!\perp y$，则 $x/(x+y) \sim \mathrm{Beta}(n/2,\, m/2)$。）

**Solution to Exercise 1.1.** **(a)** $x$ 的边际密度，对 $|t| \le 1$：

$$p(t) = \int_{-\infty}^{\infty} f(t,y)\,\mathrm{d}y = \frac{1}{\pi}\int_{-\sqrt{1-t^2}}^{\sqrt{1-t^2}} \mathrm{d}y = \frac{2\sqrt{1-t^2}}{\pi}.$$

这正是半圆律。直觉上，在 $x = t$ 处切一刀，截得弦长为 $2\sqrt{1-t^2}$，除以总面积 $\pi$ 就得到密度。

**(b)** 令 $u = (x+1)/2$，则 $x = 2u - 1$， $\mathrm{d}x = 2\,\mathrm{d}u$，$u \in [0,1]$。代入半圆律密度：

$$p_U(u) = p(2u-1)\cdot 2 = \frac{2}{\pi}\sqrt{1-(2u-1)^2}\cdot 2 = \frac{4}{\pi}\sqrt{4u(1-u)} = \frac{8}{\pi}\,u^{1/2}(1-u)^{1/2}.$$

而 $\mathrm{Beta}(3/2, 3/2)$ 的密度为 $\frac{\Gamma(3)}{\Gamma(3/2)^2}\,u^{1/2}(1-u)^{1/2}$。因为 $\Gamma(3) = 2$，$\Gamma(3/2) = \frac{\sqrt{\pi}}{2}$，所以

$$\frac{\Gamma(3)}{\Gamma(3/2)^2} = \frac{2}{\pi/4} = \frac{8}{\pi}.$$

两者一致，故 $(x+1)/2 \sim \mathrm{Beta}(3/2, 3/2)$。

**(c)** 令 $v = x^2$，则 $v \in [0,1]$。因为 $x$ 的密度关于 0 对称，$v$ 的分布函数为

$$P(v \le t) = P(x^2 \le t) = P(-\sqrt{t} \le x \le \sqrt{t}) = 2\int_0^{\sqrt{t}} \frac{2\sqrt{1-s^2}}{\pi}\,\mathrm{d}s.$$

对 $t$ 求导：

$$p_V(t) = 2\cdot\frac{2\sqrt{1-t}}{\pi}\cdot\frac{1}{2\sqrt{t}} = \frac{2}{\pi}\,t^{-1/2}(1-t)^{1/2}, \quad 0 < t < 1.$$

Beta$(1/2,3/2)$ 的密度为 $\frac{\Gamma(2)}{\Gamma(1/2)\Gamma(3/2)}\, t^{-1/2}(1-t)^{1/2}$。代入 $\Gamma(2)=1$, $\Gamma(1/2)=\sqrt{\pi}$, $\Gamma(3/2)=\frac{\sqrt{\pi}}{2}$：完全一致

**(d)** 由 (c) 知 $x^2 \sim \mathrm{Beta}(1/2,3/2)$。利用提示中给出的性质：若 $w \sim \mathrm{Beta}(n/2,m/2)$，则 $w \overset{d}{=} \frac{\xi}{\xi+\eta}$，其中 $\xi \sim \chi^2_n$，$\eta \sim \chi^2_m$ 且独立。这里 $x^2 \sim \mathrm{Beta}(1/2,3/2)$ 意味着

$$x^2 \overset{d}{=} \frac{\xi}{\xi+\eta}, \quad \xi \sim \chi^2_1,\ \eta \sim \chi^2_3,\ \xi \perp\!\!\!\perp \eta.$$

于是

$$\frac{x^2}{1-x^2} \overset{d}{=} \frac{\xi}{\eta}.$$

因此

$$\frac{\sqrt{3}\,x}{\sqrt{1-x^2}} \overset{d}{=} \pm\sqrt{\frac{3\xi}{\eta}} = \pm\sqrt{\frac{\xi/1}{\eta/3}}.$$

注意 $\frac{\xi/1}{\eta/3} \sim F_{1,3}$，而 $t_3^2 \sim F_{1,3}$，所以 $\pm\sqrt{F_{1,3}} \sim t_3$（符号由 $x$ 的对称性保证）。综合起来，$\frac{\sqrt{3}\,x}{\sqrt{1-x^2}} \sim t_3$。

**Exercise 1.2.** 假设二元随机向量 $\mathbf{x} = (x_1,x_2)^\top \in \mathbb{R}^2$ 的概率密度函数 $f(\mathbf{x})$ 仅与 $\|\mathbf{x}\| = \sqrt{x_1^2+x_2^2}$ 有关，即存在某个一元函数 $h$ 使得 $f(\mathbf{x}) = h(\|\mathbf{x}\|) = h\Big(\sqrt{x_1^2+x_2^2}\Big)$，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

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

记 $r$ 的概率密度函数为 $p(r)$。

(a) 试证明 $\theta \sim U(0,2\pi)$，$\theta \perp\!\!\!\perp r$，且 $r$ 的概率密度 $p(r) = 2\pi r\, h(r)$。假设已知模长分布 $p(r)$，则 $h(r) = p(r)/(2\pi r)$，据此我们可得到 $\mathbf{x}$ 的概率密度

$$f(\mathbf{x}) = h(\|\mathbf{x}\|) = p(\|\mathbf{x}\|)/(2\pi\|\mathbf{x}\|).$$

基于该结论求解下述问题：

(b) 若 $r \sim \mathrm{Beta}(2,1)$，即概率密度为 $p(r) = 2r\mathbf{1}_{(0,1)}(r)$，证明 $\mathbf{x} \sim U(B_2)$。

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

(d) (Box-Muller transform) 假设 $r \sim \chi_2$ 分布（也称作 Rayleigh 分布），即其概率密度函数为

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

证明 $\mathbf{x} \sim N_2(\mathbf{0}, I_2)$，即 $x_1,x_2 \overset{\text{iid}}{\sim} N(0,1)$。（注：Box-Muller transform 是标准正态随机数最常用的产生方法。）

**Solution to Exercise 1.2.** **(a)** 做极坐标变换 $x_1 = r\cos\theta$，$x_2 = r\sin\theta$，Jacobian 行列式为

$$\left|\frac{\partial(x_1,x_2)}{\partial(r,\theta)}\right| = r.$$

$(r,\theta)$ 的联合密度为

$$f_{R,\Theta}(r,\theta) = h(r)\cdot r, \quad r>0,\ 0\le\theta<2\pi.$$

密度可以分解为只与 $r$ 有关的部分 $rh(r)$ 和只与 $\theta$ 有关的常数 1，因此 $r$ 和 $\theta$ 独立。对 $\theta$ 积分：

$$p(r) = \int_0^{2\pi} rh(r)\,\mathrm{d}\theta = 2\pi rh(r).$$

对 $r$ 积分，由归一化条件：

$$f_\Theta(\theta) = \frac{1}{2\pi}, \quad 0 \le \theta < 2\pi.$$

所以 $\theta \sim U(0, 2\pi)$。由公式 $h(r) = p(r)/(2\pi r)$，可以从模长分布反推联合密度。

**(b)** Beta(2, 1) 的密度为 $p(r) = 2r\mathbf{1}_{(0,1)}(r)$。由 (a) 的公式：

$$f(\mathbf{x}) = \frac{p(\|\mathbf{x}\|)}{2\pi\|\mathbf{x}\|} = \frac{2\|\mathbf{x}\|}{2\pi\|\mathbf{x}\|} = \frac{1}{\pi}, \quad \|\mathbf{x}\| < 1.$$

这正是 $U(B_2)$ 的密度。

**(c)** $p(r) = 1$，$0 < r < 1$。代入：

$$f(\mathbf{x}) = \frac{1}{2\pi\|\mathbf{x}\|}, \quad 0 < \|\mathbf{x}\| < 1.$$

球心处 $\|\mathbf{x}\| \to 0$ 时密度趋于 $+\infty$。直觉上，模长均匀分布意味着靠近原点的点”太多了”（小半径处圆周短，但概率和大半径一样），导致球心附近密度爆炸。

**(d)** $r$ 的密度为 $p(r) = re^{-r^2/2}$，$r > 0$（Rayleigh 分布）。代入：

$$f(\mathbf{x}) = \frac{p(\|\mathbf{x}\|)}{2\pi\|\mathbf{x}\|} = \frac{\|\mathbf{x}\|e^{-\|\mathbf{x}\|^2/2}}{2\pi\|\mathbf{x}\|} = \frac{1}{2\pi}e^{-\|\mathbf{x}\|^2/2} = \frac{1}{2\pi}e^{-(x_1^2+x_2^2)/2}.$$

这可以分解为

$$f(\mathbf{x}) = \left(\frac{1}{\sqrt{2\pi}}e^{-x_1^2/2}\right)\left(\frac{1}{\sqrt{2\pi}}e^{-x_2^2/2}\right),$$

所以 $x_1, x_2 \overset{\text{iid}}{\sim} N(0, 1)$，即 $\mathbf{x} \sim N_2(\mathbf{0}, I_2)$。这就是 Box-Muller 方法的数学基础：生成 $r$ 服从 Rayleigh 分布（等价于 $r^2 \sim \mathrm{Exp}(1/2)$，即 $r = \sqrt{-2\ln U}$，$U \sim U(0,1)$），$\theta \sim U(0, 2\pi)$，然后令 $x_1 = r\cos\theta$，$x_2 = r\sin\theta$ 就得到两个独立标准正态。

**Exercise 1.3.** 假设随机向量 $\mathbf{x} = (x_1, \ldots, x_n)^\top \sim U(B_n)$（即 $\mathbf{x}$ 服从 $\mathbb{R}^n$ 空间中单位球内的均匀分布），令 $\mathbf{x}_1 = (x_1, \ldots, x_k)^\top$，$\mathbf{x}_2 = (x_{k+1}, \ldots, x_n)^\top$，$1 \le k \le n$。

(a) 证明 $\mathbf{x}_1$ 的边际概率密度为

$$p(\mathbf{x}_1) = \frac{\Gamma\left(\frac{n}{2}+1\right)}{\pi^{k/2}\,\Gamma\left(\frac{n-k}{2}+1\right)}(1 - \|\mathbf{x}_1\|^2)^{(n-k)/2}, \quad \mathbf{x}_1 \in B_k.$$

特别地，当 $k = n - 1$ 时，$\mathbf{x}_1 = (x_1, \ldots, x_{n-1})^\top$ 的边际概率密度为”半球分布”：

$$p(\mathbf{x}_1) = \frac{2}{|B_n|}\sqrt{1 - \|\mathbf{x}_1\|^2}, \quad \mathbf{x}_1 \in B_{n-1}.$$

(b) 证明对任何 $1 \le k \le n-1$，$\mathbf{x}_1$ 不可能服从均匀分布，但 $\mathbf{x}_2|\mathbf{x}_1 \sim U(B_{n-k}(r))$，$r = \sqrt{1 - \|\mathbf{x}_1\|^2}$。

(c) 证明 $\mathbf{x}_2/(1 - \|\mathbf{x}_1\|^2)^{1/2} \sim U(B_{n-k})$ 且与 $\mathbf{x}_1$ 独立。

**Solution to Exercise 1.3.**

**(a)** $\mathbf{x} \sim U(B_n)$ 的密度为 $f(\mathbf{x}) = \frac{1}{|B_n|} = \frac{\Gamma(n/2+1)}{\pi^{n/2}}$。$\mathbf{x}_1 = (x_1, \ldots, x_k)^\top$ 的边际密度需要对 $\mathbf{x}_2 = (x_{k+1}, \ldots, x_n)^\top$ 在区域 $\|\mathbf{x}_2\|^2 \le 1 - \|\mathbf{x}_1\|^2$ 上积分:

$$p(\mathbf{x}_1) = \frac{1}{|B_n|} \cdot \left| B_{n-k}\Big(\sqrt{1-\|\mathbf{x}_1\|^2}\Big) \right| = \frac{1}{|B_n|} \cdot |B_{n-k}| \cdot (1-\|\mathbf{x}_1\|^2)^{(n-k)/2}, \quad \mathbf{x}_1 \in B_k.$$

这里用到了 $n-k$ 维球的体积公式 $|B_{n-k}(r)| = |B_{n-k}| \cdot r^{n-k}$。代入 $|B_m| = \frac{\pi^{m/2}}{\Gamma(m/2+1)}$:

$$\begin{aligned} p(\mathbf{x}_1) &= \frac{\Gamma(n/2+1)}{\pi^{n/2}} \cdot \frac{\pi^{(n-k)/2}}{\Gamma((n-k)/2+1)} \cdot (1-\|\mathbf{x}_1\|^2)^{(n-k)/2} \\ &= \frac{\Gamma(n/2+1)}{\pi^{k/2}\,\Gamma((n-k)/2+1)} (1-\|\mathbf{x}_1\|^2)^{(n-k)/2}. \end{aligned}$$

当 $k = n-1$ 时，$\Gamma(3/2) = \frac{\sqrt{\pi}}{2}$，代入得

$$p(\mathbf{x}_1) = \frac{\Gamma(n/2+1)}{\pi^{(n-1)/2} \cdot \frac{\sqrt{\pi}}{2}} \sqrt{1-\|\mathbf{x}_1\|^2} = \frac{2\,\Gamma(n/2+1)}{\pi^{n/2}} \sqrt{1-\|\mathbf{x}_1\|^2} = \frac{2}{|B_n|} \sqrt{1-\|\mathbf{x}_1\|^2}.$$

**(b)** 由 (a) 的结果，$p(\mathbf{x}_1)$ 含有因子 $(1-\|\mathbf{x}_1\|^2)^{(n-k)/2}$，这不是常数（因为 $1 \le k \le n-1$ 时 $(n-k)/2 > 0$），所以 $\mathbf{x}_1$ 不可能是 $B_k$ 上的均匀分布。

对于条件分布，令 $r = \sqrt{1-\|\mathbf{x}_1\|^2}$:

$$f(\mathbf{x}_2|\mathbf{x}_1) = \frac{f(\mathbf{x})}{p(\mathbf{x}_1)} = \frac{1/|B_n|}{p(\mathbf{x}_1)} = \frac{1}{|B_{n-k}| \cdot r^{n-k}}, \quad \|\mathbf{x}_2\| \le r.$$

这正是半径为 $r$ 的 $n-k$ 维球上均匀分布的密度，即 $\mathbf{x}_2|\mathbf{x}_1 \sim U(B_{n-k}(r))$。

**(c)** 由 (b)，给定 $\mathbf{x}_1$ 后，$\mathbf{x}_2$ 在半径 $r = \sqrt{1-\|\mathbf{x}_1\|^2}$ 的球上均匀分布。做缩放 $\mathbf{z} = \mathbf{x}_2/r = \mathbf{x}_2/(1-\|\mathbf{x}_1\|^2)^{1/2}$，则 $\mathbf{z}$ 在 $B_{n-k}$ 上均匀分布，且这个条件分布不依赖于 $\mathbf{x}_1$（给定 $\mathbf{x}_1$ 后 $\mathbf{z}$ 的密度是 $1/|B_{n-k}|$，与 $\mathbf{x}_1$ 无关）。由此 $\mathbf{z} \perp\!\!\!\perp \mathbf{x}_1$ 且 $\mathbf{z} \sim U(B_{n-k})$。

**Remark 1.1**. 球面均匀分布：$\mathbb{R}^n$ 中 $n-1$ 维单位球面 $S^{n-1} = \{\mathbf{u} \in \mathbb{R}^n : \|\mathbf{u}\| = 1\}$ 作为 $\mathbb{R}^n$ 的 $n-1$ 维子集，其 Lebesgue 测度为 0。我们称随机向量 $\mathbf{u} \in \mathbb{R}^n$ 服从 $S^{n-1}$ 上的均匀分布，记作 $\mathbf{u} \sim U(S^{n-1})$，如果

$$P(\mathbf{u} \in S_0) = |S_0|/|S^{n-1}|, \quad \forall\, S_0 \subset S^{n-1},$$

其中 $|S_0|$ 为 $S_0 \subset S^{n-1}$ 的大小 (size) 或" 面积"。$U(S^{n-1})$ 在 $\mathbb{R}^n$ 中没有概率密度。当 $n = 1$ 时，$U(S^0)$ 是离散分布（参见第 4(b) 题）；当 $n > 1$ 时，$U(S^{n-1})$ 的分布函数是连续的，我们称之为奇异连续分布（参见第 5 题）。

**Exercise 1.4.** $\mathbb{R}^1$ 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 $P(u = \pm 1) = 1/2$（即 $u \sim U(S^0)$，球面均匀分布）。

**Solution to Exercise 1.4.** **(a)** $x \sim U(-1,1)$，密度为 $\frac{1}{2}$。对 $0 < t < 1$：

$$P(|x| \le t) = P(-t \le x \le t) = t.$$

这正是 $U(0,1)$ 的分布函数。求导得 $p_r(t) = 1$。

**(b)** $P(x > 0) = P(x < 0) = \frac{1}{2}$（$x \sim U(-1,1)$），而 $P(x = 0) = 0$，所以

$$P(u = 1) = P(x > 0) = \frac{1}{2}, \quad P(u = -1) = P(x < 0) = \frac{1}{2}.$$

即 $u$ 以等概率取 $\pm 1$，这就是 $U(S^0)$。

**Exercise 1.5.** 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），即对任何圆弧 $L \subset S^1$，

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

其中 $|L|$ 为弧长（该分布在 $\mathbb{R}^2$ 中没有概率密度）。试求 $x$ 的边际分布。

**Solution to Exercise 1.5.** 参数化：$x = \cos\theta$，$y = \sin\theta$，$\theta \sim U(0, 2\pi)$。令 $t = \cos\theta$，$|t| \le 1$。对 $-1 < t < 1$，$P(x \le t) = P(\cos\theta \le t)$。$\cos\theta \le t$ 等价于 $\theta \in [\arccos t, 2\pi - \arccos t]$（在 $[0, 2\pi)$ 上），所以

$$P(x \le t) = \frac{2\pi - 2\arccos t}{2\pi} = 1 - \frac{\arccos t}{\pi}.$$

对 $t$ 求导：

$$p_X(t) = \frac{1}{\pi\sqrt{1 - t^2}}, \quad |t| < 1.$$

这是反正弦分布 (arcsine distribution)，也就是 $\mathrm{Beta}(1/2, 1/2)$ 做线性变换到 $[-1, 1]$ 后的结果。

和第 1 题的半圆律比较：一个是圆盘上的投影（中间密集），一个是圆周上的投影（两端密集），是完全不同的。

**Exercise 1.6.** （$\mathbb{R}^1$ 中的球对称分布/0 对称分布）假设 $x$ 是关于 0 对称的随机变量，$P(x = 0) = 0$。

(a) 令 $u = x/|x|$，证明 $P(u = \pm 1) = 1/2$，$u \perp\!\!\!\perp |x|$。

(b) 假设 $\sigma \perp\!\!\!\perp x$，$\sigma \sim U(S^0)$ 即 $P(\sigma = \pm 1) = 1/2$，证明 $x \overset{d}{=} x\sigma \overset{d}{=} |x|\sigma$。

**Solution to Exercise 1.6.** **(a)** $x$ 关于 0 对称意味着 $x \overset{d}{=} -x$。$u = x/|x| = \mathrm{sgn}(x)$。由对称性 $P(x > 0) = P(x < 0)$，加上 $P(x = 0) = 0$，所以 $P(u = 1) = P(u = -1) = \frac{1}{2}$。

独立性：对任何 $r > 0$，

$$P(u = 1, |x| \le r) = P(0 < x \le r) = P(-r \le x < 0) = P(u = -1, |x| \le r),$$

因此

$$P(u = 1, |x| \le r) = \frac{1}{2} P(|x| \le r) = P(u = 1) \cdot P(|x| \le r).$$

$u = -1$ 的情况完全类似。所以 $u \perp\!\!\!\perp |x|$。

**(b)** $\sigma \perp\!\!\!\perp x$，$P(\sigma = \pm 1) = \frac{1}{2}$。

$x \overset{d}{=} |x|\sigma$：由 (a)，$x = |x| \cdot u$，且 $u \perp\!\!\!\perp |x|$，$u$ 以等概率取 $\pm 1$。而 $\sigma$ 独立于 $|x|$ 且分布和 $u$ 一样，所以 $|x| \cdot u \overset{d}{=} |x| \cdot \sigma$。

$x \overset{d}{=} x\sigma$：因为 $x$ 关于 0 对称，$x \overset{d}{=} -x$。条件在 $\sigma = 1$ 时 $x\sigma = x$，$\sigma = -1$ 时 $x\sigma = -x \overset{d}{=} x$。用全概率公式，$x\sigma$ 的分布就是 $x$ 的分布。