

课程主页: <http://staff.ustc.edu.cn/~ynyang/vector>

# 第三讲 球对称分布II

2024.3.4

Bertrand Paradox

我们以几何概率问题 “Bertrand Paradox” 为例，回顾与单位圆有关的均匀性问题（以及球对称分布与径向分布的关系）

Bertrand  
Paradox

例5. 随机选取单位圆的一条弦，其长度大于圆内接等边三角形边长 $\sqrt{3}$ 的概率是多少？

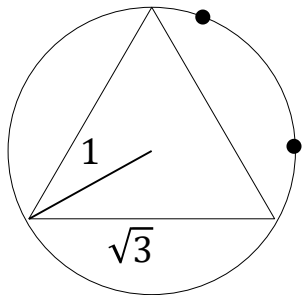
“随机选取一个点” 的含义只要指定该点（随机变量或向量）就可以确定。但“随机选取一条弦” 含义不是特别明确，这里选取的是几何体，是点的集合，不同的理解可导致三个不同的计算结果，参见：

- 苏淳,冯群强 (2020) 概率论, P29.
- [https://en.wikipedia.org/wiki/Bertrand\\_paradox\\_\(probability\)](https://en.wikipedia.org/wiki/Bertrand_paradox_(probability))。

单位圆内任意一点的极坐标表示:

$$(x_1, x_2) = (r\cos(\theta), r\sin(\theta)), 0 \leq r \leq 1, 0 \leq \theta \leq 2\pi$$

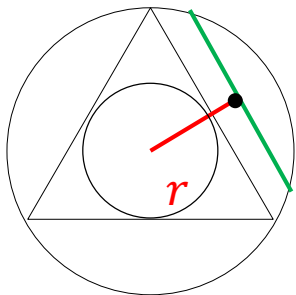
解法1: 在圆周上随机地选取两个点构成一个短弧, 若弧的长度大于 $\frac{2\pi}{3}$ , 则对应的弦长大于 $\sqrt{3}$ 。故所求概率为 $1/3$ 。



评注: 解法1将随机性局限于圆周上, 实际上假设了随机选取的点服从圆周上的均匀分布:  $(x_1, x_2) \sim U(S^1)$ ,

- $\theta \sim U(0, 2\pi)$
- $r \equiv 1$

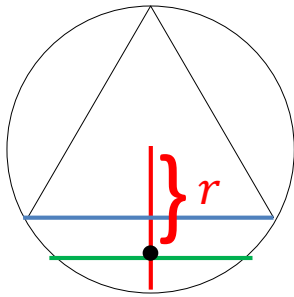
解法2: 在圆内随机地取一个点 (黑点), 垂直于径向以该点为中点得一条弦 (绿色)。这条弦如果落在三角形内切圆的内部 (半径 $1/2$ ), 则弦长大于 $\sqrt{3}$ 。内切圆面积是单位圆的 $1/4$ , 故所求概率为 $1/4$ 。



评注: 解法2将随机性分布在整个圆内, 概率值由内切圆的面积占比得到, 这实际上假设了黑点 $(x_1, x_2) \sim U(B^2)$ , 即极坐标满足:

- $\theta \sim U(0, 2\pi)$ ,
- $r = \sqrt{x_1^2 + x_2^2} \sim \sqrt{U(0, 1)}$

解法3: 在圆内随机地取一个点 (黑点), 同上得到绿色弦。旋转三角形使得弦平行于等边三角形的一条边 (蓝), 该边平分红色线段, 若黑点落在蓝线上方, 则弦长大于 $\sqrt{3}$ 。黑点落在蓝线上下方的概率都是1/2, 故所求概率为1/2



评注: 解法3假设极坐标满足

- $\theta \sim U(0, 2\pi)$ ,
- $r \sim U(0, 1)$ , 其密度  $p(r) = 1_{(0 < r < 1)}$

由此可得  $(x_1, x_2)$  的密度

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{x_1^2 + x_2^2}}, x_1^2 + x_2^2 \leq 1$$

上面  $\mathbf{x}$  的概率密度  $f(\mathbf{x})$  由  $\|\mathbf{x}\|$  的分布  $p(r)$  求得:

$$f(\mathbf{x}) = \frac{\Gamma(\frac{n}{2})}{2\pi^{\frac{n}{2}}\|\mathbf{x}\|^{n-1}} p(\|\mathbf{x}\|) \stackrel{n=1}{=} \frac{1}{2\pi\|\mathbf{x}\|} p(\|\mathbf{x}\|)$$

解法3:  $p(r) = 1_{(0 < r < 1)} \Rightarrow f(\mathbf{x}) = \frac{1}{2\pi\sqrt{x_1^2 + x_2^2}} 1_{(x_1^2 + x_2^2 \leq 1)}$ ,

解法1:  $p(r) = \delta_1(r)$ , 在1处退化  $\Rightarrow f(\mathbf{x}) = \frac{1}{2\pi} \delta_1(\|\mathbf{x}\|)$

总之，上述三种解法对于点 $(x_1, x_2) = (r\cos(\theta), r\sin(\theta))$ 的随机性都假设了方向均匀： $\theta \sim U(0, 2\pi)$ ，模长 $r$ 的随机性却有不同：

	$\theta$	$r$	$f(x_1, x_2)$
解法1	$U(0, 2\pi)$	$1_{(r=1)}$ 或 $\delta_1(r)$	$1/2\pi\delta_1(x_1^2 + x_2^2)$
解法2	$U(0, 2\pi)$	$\sqrt{U(0,1)}$	$1/\pi 1_{(x_1^2+x_2^2 \leq 1)}$
解法3	$U(0, 2\pi)$	$U(0,1)$	$1/[2\pi\sqrt{x_1^2 + x_2^2}] 1_{(x_1^2+x_2^2 \leq 1)}$

注：橙色是三种解法所做的均匀性假设，都假设了角度/方向均匀，且

解法3假设了 $r$ 均匀： $r \sim U(0,1)$ ， $p(r) = 1$ ，即Beta(1,1)

解法2假设圆盘均匀，实际上假设了 $r^2 \sim U(0,1)$ ， $p(r) = 2r$ ，即Beta(2,1)

解法1假设圆周上均匀( $r \equiv 1$ )，理解为Beta( $\infty, 1$ )

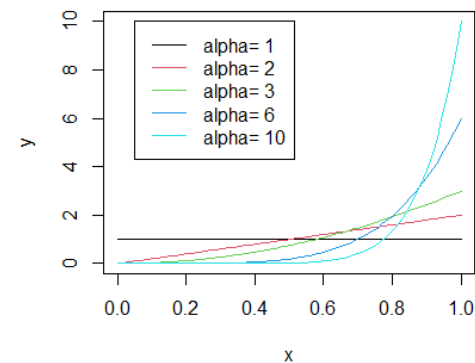
从解法3到解法1， $r = \|\mathbf{x}\|$ 在1附近的概率质量越来越大，解法1中 $r$ 在1处概率为1，（广义）密度无穷大。

Beta( $\alpha, 1$ )的密度

$$p_\alpha(r) = \alpha r^{\alpha-1}, 0 < r < 1,$$

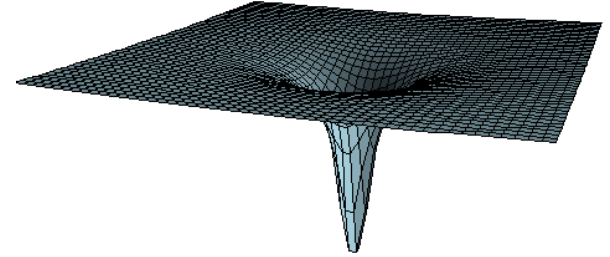
如右图， $\alpha \rightarrow \infty$ 时， $p_\alpha(r) \rightarrow \delta_1(r)$ 。

$\delta_1(r)$ : 广义函数, 在1处非0, 其它地方为0, 积分为1



$$-f(x_1, x_2) = -1/[2\pi\sqrt{x_1^2 + x_2^2}] 1_{(x_1^2+x_2^2 \leq 1)}$$

0点是奇点



至此，我们对球对称分布的半径/模长分布有了一些认识，特别地，我们介绍了球对称随机向量  $\mathbf{x}$  的分布与其模长  $\|\mathbf{x}\|$  的分布之间的互相决定的关系。

对于方向  $\mathbf{x}/\|\mathbf{x}\|$  的分布，即球面均匀分布在第一讲定理1做了介绍，下面给出证明。

第二讲定理4给出了  $\mathbf{u} \sim U(S^{n-1})$  的表示： $\mathbf{u} = \mathbf{x} / \|\mathbf{x}\|$ ， $\mathbf{x} \sim N_n(0, I_n)$ ，我们借助该表示研究球面均匀分布。为此先定义迪利克雷/迪里赫莱（Dirichlet）分布或多元beta分布。

# 狄利克雷(Dirichlete)分布

## Dirichlet 分布

定义1: 若随机向量 $\mathbf{x} = (x_1, \dots, x_n)^\top$ 的概率密度函数为

$$p(\mathbf{x}) = \frac{\Gamma(\alpha_1 + \dots + \alpha_{n+1})}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_{n+1})} \prod_{i=1}^n x_i^{\alpha_i-1} \left(1 - \sum_{i=1}^n x_i\right)^{\alpha_{n+1}-1},$$

$$0 < x_i < 1, 0 < x_1 + \dots + x_n < 1,$$

我们称 $\mathbf{x} \sim \text{Dirichlet}(\alpha_1, \dots, \alpha_n; \alpha_{n+1})$ .

分号表示地位不同

注1: beta分布是 $n = 1$ 时的Dirichlet分布:

$$p(x) = \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} x^{\alpha_1-1} (1-x)^{\alpha_2-1}, \quad 0 < x < 1,$$

记作 $x \sim \text{Beta}(\alpha_1, \alpha_2)$ 或 $x \sim \text{Dirichlet}(\alpha_1; \alpha_2)$ .

镜像对称:  $x \sim \text{Beta}(\alpha_1, \alpha_2) \Rightarrow 1-x \sim \text{Beta}(\alpha_2, \alpha_1)$ 。

注2: 定义1中若记 $x_{n+1} = 1 - \sum_{i=1}^n x_i$ , 则密度具有完美的对称形式, 因为 $\sum_{i=1}^{n+1} x_i = 1$ ,  $x_i > 0$ , 所以D分布是比率变量的分布。

注解：上页D分布的定义实际上是 $R^{n+1}$ 中单纯形（simplex）

$$\Delta^n = \{(x_1, \dots, x_{n+1})^\top : 0 < x_i < 1, x_1 + \dots + x_{n+1} = 1\}$$

上的退化分布。在单纯形上定义带约束的D分布如下（参见wiki）。

定义2：若 $(x_1, \dots, x_{n+1})^\top \in R^{n+1}$ 关于 $R^n$ 中勒贝格测度有概率密度：

$$f(x_1, \dots, x_{n+1}) = \frac{\Gamma(\sum_{i=1}^{n+1} \alpha_i)}{\prod_{i=1}^{n+1} \Gamma(\alpha_i)} \prod_{i=1}^{n+1} x_i^{\alpha_i - 1}, \quad \sum_{i=1}^{n+1} x_i = 1, \quad 0 < x_i < 1,$$

则称 $(x_1, \dots, x_{n+1})^\top \sim \text{Dirichelete}(\alpha_1, \dots, \alpha_n, \alpha_{n+1})$ 。

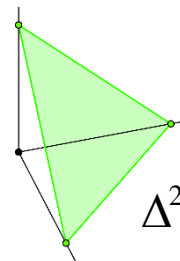
定义2在形式上有对称性，计算简单（注意这里 $\alpha_n$ 和 $\alpha_{n+1}$ 中间不再用分号），它在 $R^{n+1}$ 中没有概率密度，实际计算时我们一般还是习惯用上页的定义（有概率密度）。

如果令 $x_i = u_i^2$ ，那么

$$\Delta^n \leftrightarrow S^n = \{(u_1, \dots, u_{n+1})^\top : -1 < u_i < 1, u_1^2 + \dots + u_{n+1}^2 = 1\}$$

所以Dirichelete分布与球面上的分布有对应关系。我们考虑球面分布的时候将把球面分布映射到单纯形上的分布。按照习惯，我们在低一维的空间上（去掉和为1的约束）考虑问题，即使用定义1。

二项分布、多项分布、beta都与此类似。





命题5. 假设 $z_1, \dots, z_n, z_{n+1}$ 独立,  $z_i \sim \text{Gamma}(\alpha_i), \alpha_i > 0, i = 1, \dots, n+1$ . 令 $t = z_1 + \dots + z_n + z_{n+1}$ ,

$$x_i = z_i/t = z_i/(z_1 + \dots + z_{n+1}), i = 1, \dots, n,$$

则 $\mathbf{x} = (x_1, \dots, x_n)^\top$ 与 $t$ 独立,  $t \sim \text{Gamma}(\sum_{i=1}^{n+1} \alpha_i)$ 且 $\mathbf{x} \sim \text{Dirichlet}(\alpha_1, \dots, \alpha_n; \alpha_{n+1})$ .

证明: 容易验证Jacobi  $J = J(z_1, \dots, z_n, z_{n+1} \rightarrow x_1, \dots, x_n, t) = t^n$ .

注意 $z_i = x_i t, 1 \leq i \leq n, z_{n+1} = t - t \sum_{i=1}^n x_i$ , 所以 $x_1, \dots, x_n, t$ 的联合密度

$$\begin{aligned} p(x_1, \dots, x_n, t) &= \left( \prod_{i=1}^n \frac{1}{\Gamma(\alpha_i)} (x_i t)^{\alpha_i - 1} e^{-x_i t} \right) \frac{1}{\Gamma(\alpha_{n+1})} (t - t \sum_{i=1}^n x_i)^{\alpha_{n+1} - 1} e^{-(t - t \sum_{i=1}^n x_i)} \times J \\ &= \frac{\Gamma(\sum_{i=1}^{n+1} \alpha_i)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_{n+1})} \prod_{i=1}^n x_i^{\alpha_i - 1} \left( 1 - \sum_{i=1}^n x_i \right)^{\alpha_{n+1} - 1} \frac{1}{\Gamma(\sum_{i=1}^{n+1} \alpha_i)} t^{\alpha_1 + \cdots + \alpha_{n+1} - 1} e^{-t} \end{aligned}$$

所以 $t = z_1 + \dots + z_{n+1} \sim \text{Gamma}(\sum_{i=1}^{n+1} \alpha_i)$ ,  $t$ 与 $x_1, \dots, x_n$ 独立, 且 $\mathbf{x} = (x_1, \dots, x_n)^\top \sim \text{Dirichlet}(\alpha_1, \dots, \alpha_n; \alpha_{n+1})$ .

因为  $\chi_k^2/2 \stackrel{d}{=} \text{Gamma}(k/2)$  分布, 所以有如下推论.

推论1. 假设  $z_1, \dots, z_n, z_{n+1}$  独立,  $z_i \sim \chi_{d_i}^2, i = 1, \dots, n+1$ . 令

$$t = z_1 + \dots + z_n + z_{n+1}, x_i \sim \frac{z_i}{t} = \frac{z_i}{z_1 + \dots + z_{n+1}}, i = 1, \dots, n,$$

则  $\mathbf{x} = (x_1, \dots, x_n)^\top$  与  $t$  独立, 且  $\mathbf{x} \sim \text{Dirichlet}(d_1/2, \dots, d_n/2; d_{n+1}/2)$ .

特别地,  $n = 1$  时,  $\frac{\chi_n^2}{\chi_n^2 + \chi_m^2} \sim \text{Dirichlet}(n/2; m/2) = \text{Beta}(n/2, m/2)$   
(两个卡方 r.v. 独立)

推论2. 假设  $\mathbf{x} = (x_1, \dots, x_n)^\top \sim \text{Dirichlet}(\alpha_1, \dots, \alpha_n; \alpha_{n+1})$ , 则对任何  $1 \leq k \leq n$ ,  $(x_1, \dots, x_k)^\top \sim \text{Dirichlet}(\alpha_1, \dots, \alpha_k; \alpha_{k+1} + \dots + \alpha_{n+1})$ ,

特别地,  $x_i \sim \text{Dirichlet}(\alpha_i; \alpha - \alpha_i) = \text{Beta}(\alpha_i, \alpha - \alpha_i)$ , 其中  $\alpha = \sum_{j=1}^{n+1} \alpha_j$

证：假设  $z_1, \dots, z_n, z_{n+1}$  独立， $z_i \sim \text{Gamma}(\alpha_i), \alpha_i > 0, i = 1, \dots, n+1$ ，记

$$x_i = z_i / (z_1 + \dots + z_n + z_{n+1}), i = 1, \dots, n$$

$$\mathbf{x} = (x_1, \dots, x_n)^\top = (z_1, \dots, z_n)^\top / (z_1 + \dots + z_n + z_{n+1})$$

则

$$\mathbf{x} = (x_1, \dots, x_n)^\top \sim \text{Dirichlet}(\alpha_1, \dots, \alpha_n; \alpha_{n+1}),$$

记  $\tilde{z}_{k+1} = z_{k+1} + \dots + z_{n+1} \sim \text{Gamma}(\sum_{i=k+1}^{n+1} \alpha_i)$ ，则

$$(x_1, \dots, x_k)^\top = (z_1, \dots, z_k)^\top / (z_1 + \dots + z_k + \tilde{z}_{k+1}),$$

对  $z_1, \dots, z_k, \tilde{z}_{k+1}$  再次应用命题5，

$$(x_1, \dots, x_k)^\top = (z_1, \dots, z_k)^\top / (z_1 + \dots + z_k + \tilde{z}_{k+1}) \sim \text{Dirichlet}(\alpha_1, \dots, \alpha_k; \alpha_{k+1} + \dots + \alpha_{n+1}),$$

$$\mathbf{x} \sim \text{Dirichlet}(\alpha_1, \dots, \alpha_{n-1}; \alpha_n), \text{ 记 } \mathbf{a} = (\alpha_1, \dots, \alpha_{n-1})^\top, a = \alpha_1 + \dots + \alpha_{n-1} + \alpha_n, \\ \boldsymbol{\theta} = \mathbf{a} / a, \text{ 则 } E(\mathbf{x}) = \boldsymbol{\theta}, \text{ var}(\mathbf{x}) = (1 + a)^{-1} (\text{diag}(\boldsymbol{\theta}) - \boldsymbol{\theta}\boldsymbol{\theta}^\top),$$

# 球面均匀分布的边际分布

为了研究 $\mathbf{u} \sim U(S^{n-1})$ 的边际分布，我们可以采用两种策略

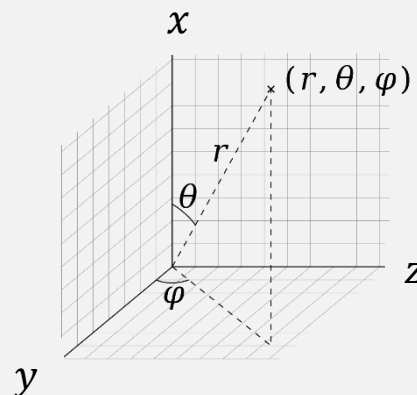
- 通过 $\mathbf{u}$ 的球坐标计算；
- 利用 $\mathbf{u}$ 的随机表示以及Dirichlete分布的性质。

我们首先以 $n = 3$ 情形为例，利用球坐标方法计算 $U(S^2)$ 的一元边际分布。更高维情况下，球坐标方法求解多元边际分布较为复杂，我们将采用第二种方法。

## 球坐标 方法

例6(参见hw2.2). 假设 $(x, y, z) \sim U(S^2)$ ，其极坐标表示(下图):

$$\begin{aligned}x &= r \cos(\theta), \\y &= r \sin(\theta) \cos(\varphi), \\z &= r \sin(\theta) \sin(\varphi), \\0 &\leq \theta \leq \pi, 0 \leq \varphi < 2\pi, \\J(x, y, z \rightarrow r, \theta, \varphi) &= r \sin(\theta). \\r &= 1\end{aligned}$$



$$q(\boldsymbol{\theta}) = \frac{\Gamma(n/2)}{2\pi^{n/2}} \sin^{n-2}(\theta_1) \sin^{n-3}(\theta_2) \cdots \sin(\theta_{n-2})$$

由第二讲定理2,  $(\theta, \varphi)$  的概率密度函数:

$$q(\theta, \varphi) = \frac{J}{|S^2|} = \frac{1}{4\pi} \sin(\theta) = \left[ \frac{1}{2} \sin(\theta) \right] \times \left[ \frac{1}{2\pi} \right]$$

所以

$$\varphi \sim U(0, 2\pi), \theta \sim f(\theta) = \frac{1}{2} \sin(\theta), \quad 0 \leq \theta \leq \pi,$$

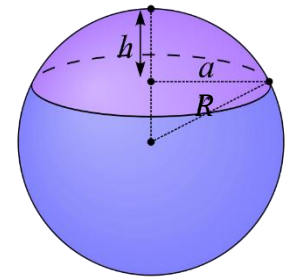
$\Rightarrow x = \cos(\theta) \sim U(-1, 1) \Rightarrow y, z \sim U(-1, 1)$ , 但  $x, y, z$  不独立。

阿基米德定理: 半径为  $R$  的三维球中, 径向高度为  $h$  的球冠的表面积  $S = 2\pi hR$  ( $S$  是  $h$  的线性函数)。

证明: 因为球面上服从均匀分布的点  $(x, y, z)$  的坐标  $x \sim U(-R, R)$ , 该点落在高度为  $h$  的球冠上的概率

$$P(R - h < x < R) = h/2R = S/(4\pi R^2) \Rightarrow S = 2\pi hR$$

这也是任何径向高度为  $h$  的球台的表面积。



二元边际  $(x, y) = (\cos(\theta), \sin(\theta)\cos(\varphi))$  的分布?

定理5（第一讲定理1）. 假设  $\mathbf{u} = (u_1, \dots, u_n)^\top \sim U(S^{n-1})$ , 则对任何  $1 \leq k \leq n-1$ ,  $\mathbf{u}_1 = (u_1, \dots, u_k)^\top$  的概率密度:

$$f_{n,k}(\mathbf{u}_1) = \frac{\Gamma(n/2)}{\pi^{k/2} \Gamma((n-k)/2)} (1 - \|\mathbf{u}_1\|^2)^{(n-k-2)/2}, \|\mathbf{u}_1\| \leq 1.$$

下面利用  $\mathbf{u} \sim U(S^{n-1})$  的随机表示证明。

大致思路如下:  $\mathbf{u} \sim U(S^{n-1})$  可以随机表示为标准正态随机向量的单位化:

$$\mathbf{u} = \mathbf{x} / \|\mathbf{x}\|, \quad \mathbf{x} \sim N_n(\mathbf{0}, I_n),$$

其“平方(诸分量平方)”

$$\mathbf{u}^2 = (u_1^2, \dots, u_n^2)^\top = \frac{\mathbf{x}^2}{\|\mathbf{x}\|^2} \sim \text{Dirichlete},$$

而Dirichlete的边际可由推论2得到。反变换（开平方）即可得到 $\mathbf{u}$ 的边际。

证明：不妨假设  $\mathbf{u} = \mathbf{x}/\|\mathbf{x}\|$ ，其中  $\mathbf{x} = (x_1, \dots, x_n)^\top \sim N_n(\mathbf{0}, I_n)$ ，

所以  $u_i = x_i/\sqrt{x_1^2 + \dots + x_n^2}$ ，令  $y_i = u_i^2 = x_i^2/(x_1^2 + \dots + x_n^2)$ ，

则  $\mathbf{y} = (y_1, \dots, y_k)^\top \sim \text{Dirichlet}\left(\frac{1}{2}, \dots, \frac{1}{2}; \frac{n-k}{2}\right)$ ，概率密度函数为

$$f(\mathbf{y}) = \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n}{2}\right)^k \Gamma\left(\frac{n-k}{2}\right)} y_1^{-1/2} \dots y_k^{-1/2} [1 - (y_1 + \dots + y_k)]^{(n-k)/2-1}$$

$\mathbf{y} = (y_1, \dots, y_k)^\top$  到  $\mathbf{u}_1 = (u_1, \dots, u_k)^\top$  的变换不是一一映射：

$$\mathbf{u}_1 = (\pm\sqrt{y_1}, \dots, \pm\sqrt{y_k})^\top,$$

因为对称性，我们仅考虑一种情况： $\mathbf{u}_1 = (\sqrt{y_1}, \dots, \sqrt{y_k})^\top$

（共  $2^k$  种），其Jacobi  $J_+(\mathbf{y} \rightarrow \mathbf{u}_1) = 2^k u_1 \dots u_k$ ，所以  $\mathbf{u}_1$  的概率密度

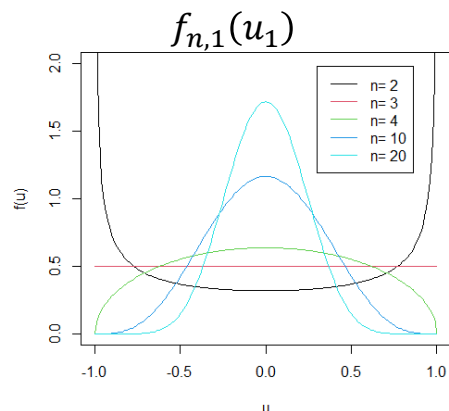
$$\begin{aligned} f_{n,k}(\mathbf{u}_1) &= f(u_1^2, \dots, u_k^2) \times J_+(\mathbf{y} \rightarrow \mathbf{u}_1)/2^k \\ &= \frac{\Gamma(n/2)}{\pi^{k/2} \Gamma((n-k)/2)} (1 - \|\mathbf{u}_1\|^2)^{(n-k-2)/2}, \|\mathbf{u}_1\| \leq 1. \end{aligned}$$

高维球面的面积大部分集中于赤道带

推论1.  $U(S^{n-1})$  的一元边际分布称为Wigner球对称分布:

$$f_{n,1}(u_1) = \frac{\Gamma(n/2)}{\pi^{1/2}\Gamma((n-1)/2)} (1 - u_1^2)^{(n-3)/2}, |u_1| < 1.$$

换言之,  $\frac{1}{2}(u_1 + 1) \sim \text{Beta}\left(\frac{n-1}{2}, \frac{n-1}{2}\right)$ ,



- 对任何  $n > 1$ ,  $E(u_1) = 0, \text{var}(u_1) = 1/n$ .
- 一元边际  $u_1$  集中于 0 附近, 由切比雪夫不等式(concentration inequality)

$$P(|u_1| < \varepsilon) > 1 - \frac{1}{n\varepsilon^2}$$

当  $n \rightarrow \infty$  时,  $P(|u_1| < \varepsilon) \rightarrow 1$ , 高维球面面积大部分在赤道带附近.

- $n = 2$  时,  $u_1 \sim f_{2,1}(u_1) = \frac{1}{\pi\sqrt{1-u_1^2}}, |u_1| < 1$  (arcsin law).
- $n = 3$  时,  $u_1 \sim f_{3,1}(u_1) = 1/2, |u_1| < 1$  (阿基米德);  $R^3$  特殊!
- 当  $n \rightarrow \infty$  时,  $\sqrt{nu_1} \xrightarrow{d} N(0, 1)$ ,  $u_1$  近似服从  $N(0, n^{-1})$

推广的阿基米德定理

推论2. (降2个维度仍然均匀)  $U(S^{n-1})$  的  $k = n - 2$  元边际分布:

$$f(\mathbf{u}_1) = \frac{\Gamma(n/2)}{\pi^{(n-2)/2}}, \quad \mathbf{u}_1 = \mathbf{u}_{[1:(n-2)]} \in B^{n-2}$$

是  $R^{n-2}$  中单位球内的均匀分布  $U(B^{n-2})$



## 球对称分布随机向量的生成

多元球对称分布正交变换不变，概率密度仅与模长有关，即模长  $\|\mathbf{x}\|$  的一元分布  $p(r)$  唯一决定了  $\mathbf{x}$  的多元球对称分布  $f(\mathbf{x})$ 。不同的一元分布  $p(r), r > 0$  (比如beta, gamma,  $\chi$ 等分布) 导致不同的球分布。

其中  $N_n(\mathbf{0}, I_n)$  分布由  $\chi_n$  (以及球面均匀分布) 生成。

但我们更为熟悉的多元标准正态随机向量的生成方式是：

$$x_1, \dots, x_n \text{ iid } \sim N(0,1) \Rightarrow (x_1, \dots, x_n)^\top \sim N_n(\mathbf{0}, I_n)$$

问题：其它球对称分布是否也可以由  $n$  个独立同分布的一元随机变量拼接得到？

Maxwell-Hershell定理说明正态是唯一的：

Maxwell - Hershell定理：

$\mathbf{x} \sim N_n(\mathbf{0}, \sigma^2 I_n) \Leftrightarrow \mathbf{x}$  服从球对称分布，且所有分量独立。

证明参见 M.Bilodeau, D.Brenner (1999) P51

# 椭球分布

假设 $\mathbf{x}$ 服从球对称分布，密度函数为 $f(\mathbf{x}) = h(\|\mathbf{x}\|)$ ,  $\mathbf{x} \in R^n$ 。给定常数向量 $\boldsymbol{\mu} \in R^n$ 和  $n$  阶正定矩阵 $\Sigma$ ，则 $\mathbf{y} = \Sigma^{1/2} \mathbf{x} + \boldsymbol{\mu}$ 的分布称为椭球分布：

$$\mathbf{y} = \Sigma^{1/2} \mathbf{x} + \boldsymbol{\mu} \sim E_p(\boldsymbol{\mu}, \Sigma)$$

其概率密度函数为

$$g(\mathbf{y}) = |\Sigma|^{-1/2} h((\mathbf{y} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{y} - \boldsymbol{\mu}))$$

$E_p(\boldsymbol{\mu}, \Sigma)$ 的均值为 $\boldsymbol{\mu}$ ，协方差为 $\Sigma$ 。

假设 $\mathbf{x}$ 服从标准的多元正态分布 $f(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}} \exp(-\|\mathbf{x}\|^2/2)$ ，则 $\mathbf{y} = \Sigma^{1/2} \mathbf{x} + \boldsymbol{\mu}$ 的概率密度函数为

$$g(\mathbf{y}) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp(-(\mathbf{y} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{y} - \boldsymbol{\mu})/2)$$

这称为多元（椭球）正态分布， $\mathbf{y} \sim N_n(\boldsymbol{\mu}, \Sigma)$ 。

同样，由标准的多元Cauchy分布  $p(\mathbf{x}) = \frac{\Gamma((n+1)/2)}{\pi^{(n+1)/2}} (1 + \|\mathbf{x}\|^2)^{-\frac{n+1}{2}}$ ,

我们可生成一般的多元(椭球) Cauchy分布：

$$g(\mathbf{y}) = \frac{\Gamma((n+1)/2)}{\pi^{(n+1)/2}} |\Sigma|^{-1/2} (1 + (\mathbf{y} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{y} - \boldsymbol{\mu}))^{-\frac{n+1}{2}}$$

球对称分布总结:

□ 多元标准正态分布是正交不变(或球对称)的

$$\mathbf{x} \sim N_n(\mathbf{0}, I_n) \Rightarrow H\mathbf{x} \sim N_n(\mathbf{0}, I_n)$$

其中 $H$ 是正交矩阵。 $\mathbf{x}$ 的分布由其模长 $r = \|\mathbf{x}\|$ 的分布即 $\chi$ 分布(根号卡方)和球面均匀分布唯一决定:

$$\mathbf{x} \stackrel{d}{=} \mathbf{u}r, \quad \mathbf{u} \sim U(S^{n-1}), \quad r \sim \chi_n, \quad r \perp \mathbf{u}$$

□ 任何球对称分布满足 $\mathbf{x} \stackrel{d}{=} H\mathbf{x}$  ( $H$ 正交), 都有表示

$$\mathbf{x} \stackrel{d}{=} \mathbf{u}r, \quad \mathbf{u} \sim U(S^{n-1}), \quad r \perp \mathbf{u}$$

它由 $r$ 的分布唯一决定(考虑到球面均匀分布的唯一性)。

□ 对球对称随机向量做仿射变换 $\mathbf{y} = A\mathbf{x} + \boldsymbol{\mu}$ , 则我们得到一大类椭球分布(包括 $N_n(\boldsymbol{\mu}, AA^T)$ )。

# 补充说明：关于同分布记号 $\stackrel{d}{=}$

(第二讲) 引理1: 若随机向量  $\mathbf{x} \stackrel{d}{=} \mathbf{y}$ , 则  $g(\mathbf{x}) \stackrel{d}{=} g(\mathbf{y})$ , 其中  $g$  是任一 (可测) 向量函数。

注: 函数  $g$  几乎没有任何限制 (但需可测), 该引理的关键是: 等号两边的随机向量 **内部做同样的运算/函数**, 分布依旧相同。

例子: 已知  $\mathbf{x} \stackrel{d}{=} \mathbf{y}$ , 何时  $\mathbf{x} + \mathbf{z} \stackrel{d}{=} \mathbf{y} + \mathbf{z}$ ?

由引理1, 一个充分条件是

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{z} \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix}$$

在两边同时应用函数  $g(\mathbf{a}, \mathbf{b}) = \mathbf{a} + \mathbf{b}$ , 得  $\mathbf{x} + \mathbf{z} \stackrel{d}{=} \mathbf{y} + \mathbf{z}$   
特别地, 当  $\mathbf{z}$  与  $\mathbf{x}, \mathbf{y}$  都独立时成立。

第二讲命题3的证明本质上应用了下述引理（参见Bilodeau&Brenner P21-22）

引理2 (Cramer-Wold定理) 假设随机向量 $\mathbf{x}, \mathbf{y} \in R^n$ ，则

$$\mathbf{x} \stackrel{d}{=} \mathbf{y} \Leftrightarrow \mathbf{t}^\top \mathbf{x} \stackrel{d}{=} \mathbf{t}^\top \mathbf{y}, \forall \mathbf{t} \in S^{n-1}.$$

证明:( $\Rightarrow$ ) 由引理1立得。

( $\Leftarrow$ ) 对任何 $\forall \mathbf{t} \in R^n, \mathbf{t}/\|\mathbf{t}\| \in S^{n-1}, \mathbf{t}^\top \mathbf{x}/\|\mathbf{t}\| \stackrel{d}{=} \mathbf{t}^\top \mathbf{y}/\|\mathbf{t}\|$

$\Rightarrow \mathbf{t}^\top \mathbf{x} \stackrel{d}{=} \mathbf{t}^\top \mathbf{y} \Rightarrow$  特征函数  $E(\exp(i \mathbf{t}^\top \mathbf{x})) = E(\exp(i \mathbf{t}^\top \mathbf{y})) \Rightarrow \mathbf{x} \stackrel{d}{=} \mathbf{y}$

注：原始的Cramer-Wold定理是关于依分布收敛的

Cramer-Wold定理：

假设  $\mathbf{x}_n, \mathbf{x} \in R^n$ ，则  $\mathbf{x}_n \xrightarrow{d} \mathbf{x} \Leftrightarrow \mathbf{t}^\top \mathbf{x}_n \xrightarrow{d} \mathbf{t}^\top \mathbf{x}, \forall \mathbf{t} \in S^{n-1}$  或  $\forall \mathbf{t} \in R^n$ .

# 补充说明： $\chi_n$ 的渐近分布

Delta方法：若 $\sqrt{n}(X_n - \theta) \xrightarrow{d} N(0, \sigma^2)$ , 则  
 $\sqrt{n}(g(X_n) - g(\theta)) \xrightarrow{d} N(0, \sigma^2 g'(\theta)^2)$ .

第二讲P9：假设 $\xi_n \sim \chi_n^2$ , 我们已知 $n \rightarrow \infty$ 时,

$$(\xi_n - n)/\sqrt{2n} \xrightarrow{d} N(0,1),$$

或近似地 $\xi_n \sim N(n, 2n)$ , 则一定有 $\sqrt{\xi_n} - \sqrt{n} \xrightarrow{d} N(0, 1/2)$

证1(delta方法): 首先将条件 $(\xi_n - n)/\sqrt{2n} \xrightarrow{d} N(0,1)$ 改写成标准形式

$$(\xi_n - n)/\sqrt{n} = \sqrt{n}(\xi_n/n - 1) \xrightarrow{d} N(0,2)$$

应用delta方法 $g(x) = \sqrt{x}$ ,  $g'(1) = 1/2$ ,

$$\sqrt{n}(g(\xi_n/n) - g(1)) = \sqrt{n}(\sqrt{\xi_n/n} - 1) \xrightarrow{d} N(0, 2 \times 1/4)$$

即 $\sqrt{\xi_n} - \sqrt{n} \xrightarrow{d} N(0, 1/2)$ .

Slusky引理:  $X_n \xrightarrow{d} X, Y_n \xrightarrow{P} c$ , 则  
 $X_n + Y_n \xrightarrow{d} X + c, X_n Y_n \xrightarrow{d} cX, X_n/Y_n \xrightarrow{d} X/c$

证2: 
$$\sqrt{\xi_n} - \sqrt{n} = \frac{\xi_n - n}{\sqrt{\xi_n + \sqrt{n}}} = \frac{(\xi_n - n)/\sqrt{n}}{\sqrt{\xi_n/n + 1}}$$

因为  $\frac{\xi_n - n}{\sqrt{n}} \xrightarrow{d} N(0,2)$ , 而  $\xi_n$  可表示为  $n$  个独立标准正态随机变量平方和, 由大数律  $\xi_n/n \xrightarrow{P} 1$ ,

由Slusky引理,  $\sqrt{\xi_n} - \sqrt{n} \xrightarrow{d} N(0,1/2)$ .