A method to construct exponential families by representation theory

Koichi Tojo · Taro Yoshino

Received: 13 May 2022 / Revised: 30 September 2022 / Accepted: 8 October 2022 / Published online: 23 October 2022
© The Author(s), under exclusive licence to Springer Nature Singapore Pte Ltd. 2022

Abstract
In this paper, we give a method to construct “good” exponential families systematically by representation theory. More precisely, we consider a homogeneous space \( G/H \) as a sample space and construct an exponential family invariant under the transformation group \( G \) by using a representation of \( G \). The method generates widely used exponential families such as normal, gamma, Bernoulli, categorical, Wishart, von Mises, Fisher–Bingham and hyperboloid distributions.

Keywords
Exponential family · Representation theory · Homogeneous space · Transformation model · Harmonic exponential family

Mathematics Subject Classification
62H10 · 62H11 · 20G05 · 22F30

Contents
1 Introduction ............................................. 494
2 Related work ............................................. 494
  2.1 Transformation model ...................................... 494
  2.2 Harmonic exponential families ................................. 494
3 A method to construct good exponential families .......................... 495
  3.1 Our method ........................................... 495

Communicated by Hiroshi Matsuzoe.

Koichi Tojo
koichi.tojo@riken.jp

Taro Yoshino
yoshino@ms.u-tokyo.ac.jp

1 RIKEN Center for Advanced Intelligence Project, Nihonbashi 1-chome Mitsui Building, 15th Floor, 1-4-1 Nihonbashi, Chuo-ku, Tokyo 103-0027, Japan

2 Graduate School of Mathematical Science, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8914, Japan
1 Introduction

Exponential families are important subjects of study in the field of information geometry and are often used for Bayesian inference because they have conjugate priors, which have been well studied [7]. By definition, there are infinitely many exponential families, however, only a small part of them are widely used such as Bernoulli, categorical, normal, multivariate normal, gamma, inverse gamma, von Mises, von Mises–Fisher, Fisher–Bingham, Wishart and hyperboloid distributions, which have been investigated individually.

We want to understand them systematically. It should be nice if we have a method to construct exponential families which has the following two properties:

(i) The method can generate many well-known good distributions.

(ii) Distributions obtained by the method are limited in principle.

In this paper, we introduce a method satisfying them. In our method, representation theory plays an important role. Actually, most of useful distributions have the same symmetry as the sample spaces. More precisely, the sample space can be regarded as a homogeneous space $G/H$ and the family of distributions is $G$-invariant.

In Sect. 3, we introduce our method and show that the obtained families are invariant under the transformation groups $G$. In Sect. 4, by the method, we realize examples as mentioned above.

2 Related work

2.1 Transformation model

A transformation model is one of the parametric models in statistics, which enables us to understand many families of probability measures on sample spaces comprehensively and have good properties on statistics [4, 5]. However, transformation model has a problem that the overlap between transformation models and exponential families is too small. Therefore, it fails to include many important families of probability measures on sample spaces which admit group actions (see Section 1.3 in [2]). Our method clear up the problem by using representation theory. In fact, our method gen-
erates many useful exponential families invariant under a transformation group (see Table 1).

2.2 Harmonic exponential families

Our method can be regarded as a generalization of harmonic exponential families suggested in [6]. They studied densities on a compact Lie group $G$ or its homogeneous spaces $G/H$ such as a torus or a sphere by using unitary representations of $G$. As mentioned in Section 7 (“Discussion and Future Work”) in [6], they wished an extension to noncompact groups. Our method results in a generalization of their method including noncompact groups. In the case where $G$ is compact, there is a unique invariant measure on $G$ or $G/H$ up to positive constant. On the other hand, in the case where $G$ is noncompact, it does not hold. For example, let $G$ be the affine transformation group $\mathbb{R}^\times \ltimes \mathbb{R}$ and $H$ the closed subgroup $\mathbb{R}^\times$, then $G/H \simeq \mathbb{R}$ admits no $G$-invariant measures. However, $G/H$ admits relatively $G$-invariant measures. Therefore, we use a relatively $G$-invariant measures on $G/H$ instead of a $G$-invariant measure. Since a relatively $G$-invariant measure on $G/H$ is not unique, we have a lot of choice of relatively $G$-invariant measures. Our method does not choose one of them but take “all” of them (see Sect. 3.1).

3 A method to construct good exponential families

In this section, we introduce a method to construct a family of distributions on a homogeneous space $G/H$ by representation theory. Here $G$ is a Lie group with finitely many connected components and $H$ is a closed subgroup of $G$. Moreover, we show that the families obtained by our method are exponential families on $G/H$.

3.1 Our method

We put $X := G/H$ as above, which is regarded as a sample space. Our method generates a family $\{p_\theta\}_{\theta \in \Theta}$ of probability measures on $X$ from two inputs as follows:

(i) Let $\rho : G \rightarrow GL(V)$ be a representation of $G$ on a finite dimensional real vector space $V$.
(ii) Let $v_0$ be an $H$-fixed vector in $V$.

We consider the following vector space (see Fact 3.5): $W_0(G, H) := \{\tau : G \rightarrow \mathbb{R} : \tau \text{ is a continuous group homomorphism with } \tau|_H = 0\}$.

We assume that there exist nonzero relatively $G$-invariant measures on $G/H$ and take one of them, say $\nu$. The family $\mathcal{P}$ below does not depend on the choice of $\nu$ (Remark 3.1). Then we set
\[ d\tilde{p}_\theta(x) := d\tilde{p}_{\varphi, \tau}(x) = \exp(-\langle \varphi, \rho(x)v_0 \rangle + \tau(x))d\nu(x) \]

\[
\Theta := \left\{ \theta = (\varphi, \tau) \in V^\vee \times W_0(G, H), x = gH \in G/H, \int_X d\tilde{p}_\theta < \infty \right\},
\]

\[ c_\theta := \left( \int_X d\tilde{p}_\theta \right)^{-1} (\theta \in \Theta). \]

Here we denote the dual space of \( V \) by \( V^\vee \).

From the conditions that \( v_0 \) is \( H \)-fixed and \( \tau|_H = 0 \), the notations \( \rho(x)v_0 \) and \( \tau(x) \) are well-defined. Then we get a probability measure \( p_\theta \) on \( X \) as follows:

\[ p_\theta := c_\theta \tilde{p}_\theta (\theta \in \Theta). \]

As a result, we obtain the following family of distributions on \( X \):

\[ \mathcal{P} := \{ p_\theta \}_{\theta \in \Theta}. \]

**Remark 3.1** This family \( \mathcal{P} \) does not depend on the choice of \( \nu \). In fact, for another nonzero relatively \( G \)-invariant measure \( \nu' \) on \( G/H \), there exist \( c \in \mathbb{R}_{>0} \) and \( \tau \in W_0(G, H) \) such that \( d\nu'(x) = ce^\tau(x)d\nu(x) \) by the next fact.

**Fact 3.2** (See [13, Corollary 4.27] for example) Assume that there exist nonzero relatively \( G \)-invariant measures on \( G/H \), and take one of them, say \( \nu_0 \). Then, the set of all nonzero relatively \( G \)-invariant measures on \( G/H \) is given as follows:

\[ \{ ce^\tau \nu_0 | c \in \mathbb{R}_{>0}, \tau \in W_0(G, H) \}. \]

### 3.2 Exponential family

In this subsection, we see that families obtained by our method are exponential families. There are several definitions of an exponential family. We adopt the following Definition 3.3 and show that families of distributions obtained by our method are exponential families in this sense.

Let \( X \) be a manifold and \( \mathcal{R}(X) \) the set of all Radon measures on \( X \).

**Definition 3.3** (Exponential family, see Section 5 in [1] for example) Let \( \mathcal{P} \) be a nonempty subset of \( \mathcal{R}(X) \) consisting of probability measures. \( \mathcal{P} \) is called an exponential family on \( X \) if there exists a triple \( (\mu, V, T) \) satisfying the following four conditions:

(i) \( \mu \in \mathcal{R}(X) \),

(ii) \( V \) is a finite dimensional real vector space,

(iii) \( T : X \to V \) is a continuous map,

(iv) for any \( p \in \mathcal{P} \), there exists \( \theta \in V^\vee \) such that

\[ dp(x) = c_\theta e^{-\langle \theta, T(x) \rangle}d\mu(x), \]
where $V^\vee$ is the dual space of $V$, and

$$c_\theta := \left( \int_{x \in X} e^{-\langle \theta, T(x) \rangle} d\mu(x) \right)^{-1}.$$  

**Proposition 3.4** A family $\mathcal{P} := \{p_\theta\}_{\theta \in \Theta}$ obtained by the method in Sect. 3.1 is an exponential family on $X$ if $\Theta$ is not empty.

To show this proposition, we need:

**Fact 3.5** [13, Proposition 3.3] Let $G$ be a Lie group with finitely many connected components. Then the following set is a finite dimensional vector subspace of $C(G)$:

$$\{ f \in C(G) \mid f : G \to \mathbb{R} \text{ is a group homomorphism} \}.$$  

Here, $C(G)$ denotes the set of all $\mathbb{R}$-valued continuous functions on $G$.

**Proof of Proposition 3.4** From Fact 3.5, $W_0(G, H)$ is a finite dimensional real vector space. Therefore $\tilde{V} := V \oplus W_0(G, H)^\vee$ is also a finite dimensional vector space. Then we define $T : X \to \tilde{V}$ by

$$T(x) := (\rho(x)v_0, ev_x),$$  

where $ev_x : W_0(G, H) \to \mathbb{R}$ is the evaluation map at $x \in X$. We regard $\tilde{V}$ as $V$ of the condition (ii) in Definition 3.3. Then, the other conditions in Definition 3.3 are satisfied.

### 3.3 $G$-invariance

In this subsection, we see that families obtained by our method are $G$-invariant. We put $X := G/H$, then we have a natural $G$-action on $\mathcal{R}(X)$ induced by pushforward.

**Definition 3.6** A family $\mathcal{P} \subset \mathcal{R}(X)$ is said to be $G$-invariant if

$$g \cdot p \in \mathcal{P} \quad \text{for any } p \in \mathcal{P} \text{ and } g \in G.$$  

**Proposition 3.7** The family $\mathcal{P} := \{p_\theta\}_{\theta \in \Theta}$ obtained by the method in Sect. 3.1 is $G$-invariant. Moreover, the $G$-action on $\Theta$ is given by

$$G \times \Theta \to \Theta, (g, (\varphi, \tau)) \mapsto (\rho^\vee(g)\varphi, \tau).$$  

Here, $\rho^\vee$ is the contragredient representation of $\rho$.

**Proof** Take any $\theta := (\varphi, \tau) \in \Theta$ and $g \in G$. Put $\theta' := (\rho^\vee(g)\varphi, \tau) \in V^\vee \times W_0(G, H)$. We show that $\theta' \in \Theta$ and $g \cdot p_\theta = p_{\theta'}$. We have
Table 1 Examples and inputs \((G, H, V, v_0)\) for them

| Distribution        | Sample space | \(G\) | \(H\) | \(V\) | \(v_0\) |
|---------------------|--------------|-------|-------|-------|--------|
| Bernoulli           | \([\pm 1]\)  | \([\pm 1]\) | \([1]\)  | \(\mathbb{R}_{sgn}\) | 1      |
| Categorical         | \([1, \ldots, n]\) | \(\mathcal{G}_n\) | \(\mathcal{G}_{n-1}\) | \(W\) | \(w\)  |
| Normal              | \(\mathbb{R}\) | \(\mathbb{R}^\times \times \mathbb{R}\) | \(\mathbb{R}^\times\) | \(\text{Sym}(2, \mathbb{R})\) | \(E_{22}\) |
| Multivariate normal | \(\mathbb{R}^n\) | \(\text{GL}(n, \mathbb{R}) \times \mathbb{R}^n\) | \(\text{GL}(n, \mathbb{R})\) | \(\text{Sym}(n + 1, \mathbb{R})\) | \(E_{n+1,n+1}\) |
| Gamma               | \(\mathbb{R}_{>0}\) | \(\mathbb{R}_{>0}\) | \([1]\)  | \(\mathbb{R}_1\) | 1      |
| Inverse gamma       | \(\mathbb{R}_{>0}\) | \(\mathbb{R}_{>0}\) | \([1]\)  | \(\mathbb{R}_{-1}\) | 1      |
| Wishart             | \(\text{Sym}^+(n, \mathbb{R})\) | \(\text{GL}(n, \mathbb{R})\) | \(O(n)\) | \(\text{Sym}(n, \mathbb{R})\) | \(I_n\) |
| Von Mises           | \(S^1\) | \(SO(2)\) | \(\{I_2\}\) | \(\mathbb{R}^2\) | \(e_1\) |
| Von Mises–Fisher    | \(S^{n-1}\) | \(SO(n)\) | \(SO(n - 1)\) | \(\mathbb{R}^n\) | \(e_1\) |
| Fisher–Bingham      | \(S^{n-1}\) | \(SO(n)\) | \(SO(n - 1)\) | \(\mathbb{R}^n \oplus \text{Sym}(n, \mathbb{R})\) | \(e_1 \oplus E_{11}\) |
| Hyperboloid         | \(H^n\) | \(SO_0(1, n)\) | \(SO(n)\) | \(\mathbb{R}^{n+1}\) | \(e_0\) |
| Poincaré            | \(\mathcal{H}\) | \(SL(2, \mathbb{R})\) | \(SO(2)\) | \(\text{Sym}(2, \mathbb{R})\) | \(I_2\) |

Here \(W := \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid \sum_{i=1}^n x_i = 0\}, w := \{-(n - 1), 1, \ldots, 1\} \in W\)

\[
d(g \cdot p_\theta)(x) = c_\theta \exp(-\langle \varphi, \rho(g^{-1}x)v_0 \rangle) + \tau(g^{-1}x)d(g \cdot v)(x) \\
= c_\theta \exp(-\langle \varphi, \rho(g^{-1})\rho(x)v_0 \rangle + \tau(g^{-1}) + \tau(x)) \chi(g^{-1})d\tilde{v}(x) \\
= c_\theta \exp(-\phi(g) \chi(g^{-1}) \exp(-\langle \rho'(g)\varphi, \rho(x)v_0 \rangle + \tau(x))d\tilde{v}(x) \\
= c_\theta \exp(-\tau(g)) \chi(g^{-1})d\tilde{p}_\varphi(x).
\]

Here, \(\chi : G \to \mathbb{R}_{>0}\) is the multiplier of \(v\), namely, it satisfies \(d(g \cdot v)(x) = \chi(g^{-1})d\tilde{v}(x)\) for any \(g \in G\). From the equality above, \(\theta' \in \Theta\) holds. One should check that

\[
c_\theta \exp(-\tau(g)) \chi(g^{-1}) = c_{\theta'},
\]

but in fact it is not necessary. It is enough to observe that \(p_\theta\) and \(p_{\theta'}\) are probability measures. Thus, \(\mathcal{P}\) is \(G\)-invariant. \(\square\)

4 Examples

In this section, we realize widely used exponential families in Table 1 by our method.

We use Fact 4.1 and Remark 4.2 below to determine \(W_0(G, H)\) for pairs \((G, H)\) in the following subsections.

Fact 4.1 [13, Proposition 4.28] Let \(G\) be a Lie group with finitely many connected components and \(H\) a closed subgroup of \(G\). We denote by \(\mathfrak{g}\) and \(\mathfrak{h}\) the Lie algebras of \(G\) and \(H\), respectively. Then we have \(W_0(G, H) = \{0\}\) if the pair \((G, H)\) satisfies one of the following conditions:

\(\heartsuit\) Springer
A method to construct exponential families…

(i) $g = h + [g, g]$

(ii) $G$ is compact

(iii) $G$ is semisimple.

**Remark 4.2** Let $G$ be a group and $\tau : G \to \mathbb{R}$ a group homomorphism. Then $\tau$ factors $G/[G, G]$.

### 4.1 Normal and multivariate normal distributions

In this subsection, we realize the families of normal distributions on $\mathbb{R}$ and multivariate normal distributions on $\mathbb{R}^n$ by using a representation of an affine transformation group.

**Proposition 4.3** Put $G = \mathbb{R}^\times \ltimes \mathbb{R}$ and $H = \mathbb{R}^\times \subset G$. Let $\rho : G \to GL(Sym(2, \mathbb{R}))$ be a representation given as follows:

$$\rho(g)(S) := \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} S \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \ (g = (a, b) \in \mathbb{R}^\times \ltimes \mathbb{R}, S \in Sym(2, \mathbb{R})).$$

Then $v_0 := \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \in Sym(2, \mathbb{R})$ is $H$-fixed, and by our method we obtain the family of normal distributions on $\mathbb{R}$,

$$\left\{ \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left( -\frac{(x-\mu)^2}{2\sigma^2} \right) \right\}_{(\sigma, \mu) \in \mathbb{R}_{>0} \times \mathbb{R}}. \quad (1)$$

**Proof** We identify $G/H = (\mathbb{R}^\times \ltimes \mathbb{R})/\mathbb{R}^\times$ with $\mathbb{R}$ by the following homeomorphism:

$$G/H \to \mathbb{R}, \ gH \mapsto b \quad (g = (a, b) \in G).$$

We have $W_0(G, H) = \{0\}$ from Fact 4.1(i). In fact, $g = \mathbb{R} \oplus \mathbb{R}$, $h = \mathbb{R} \oplus \{0\}$ and $[g, g] = \{0\} \oplus \mathbb{R}$. We identify $Sym(2, \mathbb{R})^\vee$ with $Sym(2, \mathbb{R})$ by taking the following inner product on $Sym(2, \mathbb{R})$:

$$\langle x, y \rangle := Tr(xy) \quad (x, y \in Sym(2, \mathbb{R})).$$

Let $dx$ denote the Lebesgue measure on $\mathbb{R}$, which is relatively $G$-invariant. Then we have

$$d \tilde{p}_{\theta_1, \theta_2, \theta_3}(x) = \exp\left( -\text{Tr}\left( \begin{pmatrix} \theta_1 \\ \theta_2 \theta_3 \end{pmatrix} \rho(g)v_0 \right) \right) dx = \exp(- (\theta_1 x^2 + 2\theta_2 x + \theta_3)) dx \quad (g = (a, x) \in \mathbb{R}^\times \ltimes \mathbb{R}).$$

The following equalities are well-known:

$$\Theta = \left\{ \begin{pmatrix} \theta_1 \\ \theta_2 \theta_3 \end{pmatrix} \in Sym(2, \mathbb{R}) \left| \int_{\mathbb{R}} d \tilde{p}_{\theta_1, \theta_2, \theta_3} < \infty \right. \right\} \quad \blacksquare$$
\[ \{ \begin{pmatrix} \theta_1 & \theta_2 \\ \theta_2 & \theta_3 \end{pmatrix} \in \text{Sym}(2, \mathbb{R}) \mid \theta_1 > 0 \}, \]

\[ c_\theta = \sqrt{\frac{\theta_1}{\pi}} \exp \left( \frac{\theta_1 \theta_3 - \theta_2^2}{\theta_1^2} \right). \]

Then we obtain the family of normal distributions (1) by a change of variables \( \mu = -\frac{\theta_2}{\theta_1} \) and \( \sigma = \frac{1}{\sqrt{2\theta_1}}. \)

By generalizing the case above, we obtain the following:

**Proposition 4.4** Let \( n \) be a positive integer. Put \( G = GL(n, \mathbb{R}) \ltimes \mathbb{R}^n, H = GL(n, \mathbb{R}) \subset G. \) Let \( \rho: G \to GL(\text{Sym}(n+1, \mathbb{R})) \) be a representation given as follows:

\[ \rho(g)(S) := \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix} S \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix}^T \quad (g = (A, b) \in GL(n, \mathbb{R}) \ltimes \mathbb{R}^n, S \in \text{Sym}(n+1, \mathbb{R})). \]

Then \( v_0 := E_{n+1,n+1} \in \text{Sym}(n+1, \mathbb{R}) \) is \( H \)-fixed, and by our method we obtain the family of multivariate normal distributions on \( \mathbb{R}^n, \)

\[ \left\{ \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \exp \left( -\frac{1}{2}(x - \mu)\Sigma^{-1}(x - \mu) \right) \right\}_{(\Sigma, \mu) \in \text{Sym}^+(n, \mathbb{R}) \times \mathbb{R}^n}, \]

where we denotes the set of all positive definite symmetric matrices of size \( n \) by \( \text{Sym}^+(n, \mathbb{R}). \)

**Proof** We identify \( G/H = (GL(n, \mathbb{R}) \ltimes \mathbb{R}^n)/GL(n, \mathbb{R}) \) with \( \mathbb{R}^n \) by the following map:

\[ G/H \to \mathbb{R}^n, (A, b)H \mapsto b. \]

We have \( W_0(G, H) = \{0\} \) from Fact 4.1(i). In fact, \( \mathfrak{g} = \mathbb{R}^n \oplus \mathbb{R}^n, \mathfrak{h} = \mathbb{R}^n \oplus \{0\} \) and \( [\mathfrak{g}, \mathfrak{g}] = \{0\} \oplus \mathbb{R}^n. \) We identify \( \text{Sym}(n+1, \mathbb{R})^\vee \) with \( \text{Sym}(n+1, \mathbb{R}) \) by taking the following inner product on \( \text{Sym}(n+1, \mathbb{R}): \)

\[ \langle x, y \rangle = \text{Tr}(xy) \quad (x, y \in \text{Sym}(n+1, \mathbb{R})). \]

Let \( dx \) denote the Lebesgue measure on \( \mathbb{R}^n, \) which is relatively \( G \)-invariant. Then we have

\[ d\tilde{p}_{\theta_1, \theta_2, \theta_3}(x) = \exp \left( -\text{Tr} \left( \begin{pmatrix} \theta_1 & \theta_2 \\ \theta_2 & \theta_3 \end{pmatrix} \begin{pmatrix} x'x & x'x \\ x'x & 1 \end{pmatrix} \right) \right) dx 
= \exp(-((x\theta_1x + 2\theta_2x + \theta_3)dx) (\theta_1 \in \text{Sym}(n, \mathbb{R}), \theta_2 \in \mathbb{R}^n, \theta_3 \in \mathbb{R}). \]
The following equalities are well-known:

\[ \Theta = \left\{ (\theta_1, \theta_2, \theta_3) \in \text{Sym}(n, \mathbb{R}) \times \mathbb{R}^n \times \mathbb{R} \mid \int_{\mathbb{R}^n} d\tilde{p}_{\theta_1, \theta_2, \theta_3} < \infty \right\} = \left\{ (\theta_1, \theta_2, \theta_3) \in \text{Sym}(n, \mathbb{R}) \times \mathbb{R}^n \times \mathbb{R} \mid \theta_1 \text{ is positive definite.} \right\}, \]

\[ c_\theta = \sqrt{\det \theta_1} \pi^n \exp \left( \frac{\det \begin{pmatrix} \theta_1 & \theta_2 \\ \theta_2 & \theta_3 \end{pmatrix}}{\det \theta_1} \right). \]

Then we obtain the family of multivariate normal distributions (2) by a change of variables \( \mu = -\theta_1^{-1} \theta_2 \in \mathbb{R}^n \) and \( \Sigma_1 = \frac{1}{2} \theta_1^{-1} \in \text{Sym}^+(n, \mathbb{R}). \)

4.2 Bernoulli and categorical distributions

In this subsection, we realize the families of Bernoulli distributions on \( \{-1, 1\} \) and categorical distributions on \( \{1, \ldots, n\} \) by using the signature representation of \( \{-1, 1\} \) and a natural representation of the symmetric group, respectively.

**Proposition 4.5** Put \( G = \{-1, 1\} \) and \( H = \{1\} \). Let \( \rho: \{-1, 1\} \rightarrow \mathbb{R}^\times \) be the signature representation. Then \( v_0 := 1 \) is \( H \)-fixed, and by our method, we obtain the family of Bernoulli distributions on \( \{-1, 1\} \),

\[ \left\{ s \frac{1+x}{2} (1-s) \frac{1-x}{2} dc(x) \right\}_{s \in (0,1)} \quad (x \in \{-1, 1\}), \quad (3) \]

where \( dc \) denotes the counting measure on \( \{-1, 1\} \).

**Proof** Since \( G \) is compact, we have \( W_0(G, H) = \{0\} \) from Fact 4.1(ii). Therefore we have

\[ d\tilde{p}_\theta(x) = e^{-\theta x} dc(x) = \begin{cases} e^{-\theta} & (x = 1) \\ e^\theta & (x = -1) \end{cases}, \]

where \( dc \) is the counting measure on \( G/H = \{-1, 1\} \), which is (relatively) \( G \)-invariant. This is integrable for any \( \theta \in \mathbb{R} \), so we have \( \Theta = \mathbb{R} \) and \( c_\theta = \frac{1}{e^{-\theta}+e^\theta} \). Therefore we get

\[ dp_\theta(x) = \begin{cases} \frac{e^{-\theta}}{e^{-\theta}+e^\theta} dc(x) & (x = 1) \\ \frac{e^\theta}{e^{-\theta}+e^\theta} dc(x) & (x = -1) \end{cases}. \]

Thus by a change of variable \( s = \frac{e^{-\theta}}{e^{-\theta}+e^\theta} \), we obtain the family of Bernoulli distributions (3) on \( \{-1, 1\} \).

By generalizing the case above, we obtain the following:
Proposition 4.6 Let $n$ be a positive number with $n \geq 2$, $G$ the symmetric group $S_n$ and $H$ the stabilizer $S_{n-1} = (S_n)_1$ of $1$ in $G$. Let $\rho : G \to GL(W)$ be a subrepresentation of the natural representation $G \to GL(n, \mathbb{R})$, where $W = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid \sum_{i=1}^n x_i = 0\}$. Then $v_0 := (-n - 1, 1, \ldots, 1) \in W$ is $H$-fixed, and by our method we obtain the family of categorical distributions on $\{1, \ldots, n\}$,

$$\{s_1^{\delta_{1,x}}s_2^{\delta_{2,x}} \cdots s_n^{\delta_{n,x}}dc(x)\}_{(s_1, \ldots, s_n) \in \{s_i > 0, \sum_{i=1}^n s_i = 1\}} \ (x \in \{1, 2, \ldots, n\}),$$

(4)

where $dc$ denotes the counting measure on $\{1, \ldots, n\}$ and $\delta_{i,x}$ is the Kronecker delta.

Proof Since $H = S_{n-1} = (S_n)_1$ is the stabilizer of the first element, $v_0 = (-n - 1, 1, \ldots, 1)$ is $H$-fixed. We have $W_0(G, H) = \{0\}$ from Fact 4.1(ii). By taking an inner product on $W$ which is the restriction of the standard inner product on $\mathbb{R}^n$, we identify $W^\vee$ with $W$. Then we have

$$d\tilde{\rho}_{a_1,\ldots,a_n}(k) = d\tilde{\rho}_{a_1,\ldots,a_n}(1, k)H = e^{-(a_1, \ldots, a_n)\rho((1, k))(-n - 1, 1, \ldots, 1)}dc(k) = e^{nak}dc(k) \ (1, k) \in S_n, k = 1, \ldots, n).$$

Since this is integrable for any $(a_1, \ldots, a_n) \in W$, we have $\Theta = \{(a_1, \ldots, a_n) \in W\}$ and $c_\emptyset = \frac{1}{\sum_{i=1}^n e^{an_i}}$. By a change of variables $s_k = e^{nak}/\sum_{i=1}^n e^{an_i}$, we obtain the family of categorical distributions (4).

4.3 Gamma and inverse gamma distributions

In this subsection, we realize the families of gamma and inverse gamma distributions on $\mathbb{R}_{>0}$ by using one dimensional representations of $\mathbb{R}_{>0}$.

Proposition 4.7 Put $G = \mathbb{R}_{>0}$ and $H = \{1\}$. For $\lambda \in \mathbb{R}$, we denote by $\rho_\lambda : G \to \mathbb{R}^\times$ a representation given by

$$\rho_\lambda(x) = x^\lambda.$$

Then $v_0 := 1 \in \mathbb{R}$ is $H$-fixed, and by our method we obtain the following family of distributions on $\mathbb{R}_{>0}$,

$$\left\{ \frac{|\lambda|}{\Gamma(k)} \left( \frac{x}{\theta} \right)^k \frac{e^{-\left( \frac{x}{\theta} \right)^\lambda}}{x} \right\}_{(k, \theta) \in \mathbb{R}_{>0} \times \mathbb{R}_{>0}},$$

(5)

Especially, we obtain the family of gamma distributions if $\lambda = 1$ and the family of inverse gamma distributions if $\lambda = -1$ as follows:

$$\left\{ \frac{1}{\Gamma(k)} \left( \frac{x}{\theta} \right)^k \frac{e^{-\frac{x}{\theta}}}{x} \right\}_{(k, \theta) \in \mathbb{R}_{>0} \times \mathbb{R}_{>0}} \ (\lambda = 1),$$

(6)
A method to construct exponential families…

\[
\left\{ \frac{1}{\Gamma(k)} \left( \frac{\theta}{x} \right)^k e^{-\theta x / x} \right\}_{(k, \theta) \in \mathbb{R}_{>0} \times \mathbb{R}_{>0}} (\lambda = -1). \tag{7}
\]

**Remark 4.8**
- By putting \( k = 1 \) in the case of \( \lambda = 1 \), we obtain the family of exponential distributions.
- By putting \( k = \frac{n}{2} \) \((n \in \mathbb{Z}_{>0})\) and \( \theta = 2 \) in the case of \( \lambda = 1 \), we obtain the family of Chi-squared distributions.
- In the case of \( \lambda = 2 \) and \( k = 1 \), we obtain the family of Rayleigh distributions.
- In the case of \( k = 1 \), we obtain the family of Weibull distributions with a shape parameter \( \lambda > 0 \).
- In the case of \( \lambda = -1 \) and \( k = \frac{1}{2} \), we obtain the family of unshifted Lévy distributions.

**Proof of Proposition 4.7** From the continuity of \( \tau \in W_0(G, H) \), we get

\[
W_0(G, H) = \left\{ \tau : x \mapsto \alpha \log x \mid \alpha \in \mathbb{R} \right\}.
\]

We identify \( \mathbb{R}^\vee \) with \( \mathbb{R} \) by taking the standard inner product on \( \mathbb{R} \). Then we have

\[
d\tilde{p}_{\alpha, \beta}(x) = \exp(-\beta x^\lambda) x^\alpha d\frac{x}{x}.
\]

Here, \( d\frac{x}{x} \) is the Haar measure on \( \mathbb{R}_{>0} \). \( \square \)

**Claim 1** We have

\[
\Theta := \left\{ (\alpha, \beta) \in \mathbb{R} \times \mathbb{R} \mid \int_{\mathbb{R}_{>0}} d\tilde{p}_{\alpha, \beta} < \infty \right\}
= \{(\alpha, \beta) \in \mathbb{R} \times \mathbb{R} \mid \beta > 0, \alpha \lambda > 0 \},
\]

\[
c^{-1}_{\alpha, \beta} := \int_{\mathbb{R}_{>0}} d\tilde{p}_{\alpha, \beta} = \frac{\Gamma(\frac{\alpha}{\lambda})}{\mid\lambda\mid \beta^{\frac{\alpha}{\lambda}}}.\]

This claim can be proved by a direct calculation, so we omit the proof. As a result, we get

\[
dp_{\alpha, \beta}(x) = \frac{\mid\lambda\mid \beta^{\frac{\alpha}{\lambda}}}{\Gamma(\frac{\alpha}{\lambda})} x^\alpha e^{-\beta x^\lambda} dx.
\]

By a change of variables \( k = \frac{\alpha}{\lambda}, \theta = \beta^{-\frac{1}{\lambda}} \), we obtain the family (5). \( \square \)

### 4.4 Wishart distributions

In this subsection, we realize the family of Wishart distributions on the set \( \text{Sym}^+(n, \mathbb{R}) \) of all positive definite symmetric matrices of size \( n \) by using a representation of \( GL(n, \mathbb{R}) \). The case \( n = 1 \) is essentially the same as the previous subsection.
Proposition 4.9 Let $n$ be a positive integer. Put $G = \text{GL}(n, \mathbb{R})$ and $H = O(n)$. Let $\rho : G \to \text{GL}(\text{Sym}(n, \mathbb{R}))$ be a representation as follows:

$$\rho(g)S = gS'g \quad (S \in \text{Sym}(n, \mathbb{R})).$$

Then $v_0 := I_n$ is $H$-fixed, and by our method we obtain the family of Wishart distributions on $X = G/H \simeq \text{Sym}^+(n, \mathbb{R})$,

$$\{ (\det y)^{\alpha} \prod_{k=1}^{n} \Gamma(\alpha - \frac{k-1}{2}) \exp(-\text{Tr}(yx))(\det x)^{\alpha - \frac{n+1}{2}} dx \}_{\{(y, \alpha) \in \text{Sym}^+(n, \mathbb{R}) \times \mathbb{R} | \alpha > \frac{n-1}{2} \}}.$$ (8)

Proof of Proposition 4.9 Claim $W_0(G, H) = \{ g \mapsto t \log |\det g| \ | t \in \mathbb{R} \}$. This claim follows from $G/[G, G] \simeq \mathbb{R}^\times$, Remark 4.2 and Fact 4.10.

We take an $G$-invariant measure $\frac{dx}{(\det x)^{\frac{n+1}{2}}}$ on $\text{Sym}^+(n, \mathbb{R})$ and identify $\text{Sym}(n, \mathbb{R})^\vee$ with $\text{Sym}(n, \mathbb{R})$ by the following inner product:

$$\langle x, y \rangle = \text{Tr}(xy) \quad (x, y \in \text{Sym}(n, \mathbb{R})).$$

Then we have

$$d\tilde{\rho}_{y, \alpha}(x) = \exp(-\text{Tr}(yx))(\det x)^{\alpha} \frac{dx}{(\det x)^{\frac{n+1}{2}}}.$$ 

Here we identify $\text{GL}(n, \mathbb{R})/O(n)$ with $\text{Sym}^+(n, \mathbb{R})$ by $gO(n) \mapsto x = gI_ntg$. To normalize the above measure, we use Fact 4.11. As a result, we obtain Wishart distribution (8). \qed

Fact 4.10 \{ $f : \mathbb{R}^\times \to \mathbb{R} \ | \ f$ is a continuous group homomorphism $\} = \{ x \mapsto \lambda \log |x| \ | \lambda \in \mathbb{R} \}.$

Fact 4.11 ([8, 12, 14], see also [9] for example) $\Theta = \{ (y, \alpha) \in \text{Sym}(n, \mathbb{R}) \times \mathbb{R} | y \in \text{Sym}^+(n, \mathbb{R}), \alpha > \frac{n-1}{2} \}.$

$$\int_{x \in \text{Sym}^+(n, \mathbb{R})} \exp(-\text{Tr}(yx))(\det x)^{\alpha} \frac{dx}{(\det x)^{\frac{n+1}{2}}} = \pi \frac{n(n-1)}{4} \prod_{k=1}^{n} \Gamma(\alpha - \frac{k-1}{2})(\det y)^{\alpha-x}. $$

4.5 Von Mises distribution

In this subsection, we realize the family of von Mises distributions on the circle $S^1$ by using the natural representation of $SO(2)$.
Proposition 4.12 Put $G = SO(2)$ and $H = \{I_2\}$. Let $\rho : G \to GL(\mathbb{R}^2)$ be the natural representation. Then $v_0 := e_1 \in \mathbb{R}^2$ is $H$-fixed, and by our method we obtain the family of von Mises distributions on $S^1$,

$$\left\{ \frac{1}{2\pi I_0(\kappa)} e^{-\kappa \cos(x-\mu)} \, dx \right\}_{(\kappa, \mu) \in \mathbb{R}_{>0} \times \mathbb{R}/2\pi \mathbb{Z}} . \tag{9}$$

Proof We use the identification $G/H = SO(2)/\{I_2\} \simeq S^1 \simeq \mathbb{R}/2\pi \mathbb{Z}$. From Fact 4.1(ii), we have $W_0(G, H) = \{0\}$. We identify $(\mathbb{R}^2)^\vee$ with $\mathbb{R}^2$ by taking the standard inner product on $\mathbb{R}^2$. Then we have

$$d\tilde{p}_{a,b}(x) = e^{-(a \cos x + b \sin x)} \, dx = e^{-\sqrt{a^2+b^2} \cos(x-\alpha)} \, dx,$$

where $dx$ is the uniform measure on $S^1$ with $\int_{S^1} dx = 2\pi$, which is the Haar measure on $SO(2)$, and $\alpha \in \mathbb{R}/2\pi \mathbb{Z}$ satisfies $\cos \alpha = \frac{a}{\sqrt{a^2+b^2}}$, $\sin \alpha = \frac{b}{\sqrt{a^2+b^2}}$. Since $S^1$ is compact, we have $\Theta = \{\theta = (a, b) \in \mathbb{R}^2\}$, and $c_0^{-1} = 2\pi I_0(\sqrt{a^2+b^2})$. Here $I_m(r)$ is the modified Bessel function of the first kind of order $m$ given as follows:

$$I_m(r) = \frac{1}{2\pi} \int_0^{2\pi} \cos(mx)e^{r \cos x} \, dx \quad (m \in \mathbb{Z}, \ r \in \mathbb{R}). \tag{10}$$

Therefore by a change of variables $\kappa = \sqrt{a^2+b^2}$, $\mu = \alpha$, we obtain the family of von Mises distributions (9).

\[\square\]

4.6 Von Mises–Fisher distribution

In this subsection, we realize the family of von Mises–Fisher distributions on the sphere $S^{n-1}$ by using the natural representation of $SO(n)$ for a positive integer $n \geq 2$.

Proposition 4.13 Put $G = SO(n)$ and $H = SO(n-1)$. Let $\rho : G \to GL(\mathbb{R}^n)$ be the natural representation. Then $v_0 := e_1 \in \mathbb{R}^n$ is $H$-fixed, and by our method we obtain the family of von Mises–Fisher distributions on $S^{n-1}$,

$$\left\{ \frac{1}{(2\pi)^{\frac{n}{2}} \|\mu\|^{1-\frac{n}{2}} I_{\frac{n}{2}-1}(\|\mu\|)} \exp(-^t \mu x) \, dx \right\}_{\mu \in \mathbb{R}^n} . \tag{11}$$

Here, $dx$ is the uniform measure on $S^{n-1}$ with $\int_{S^{n-1}} dx = \frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}$. Proof We use the identification $SO(n)/SO(n-1) \simeq S^{n-1}$, $gSO(n-1) \mapsto ge_1$. Since $G$ is compact, from Fact 4.1(ii), we obtain $W_0(G, H) = \{0\}$. We identify $(\mathbb{R}^n)^\vee$ with $\mathbb{R}^n$ by taking the standard inner product on $\mathbb{R}^n$. Then we have

$$d\tilde{p}_{\mu}(x) = \exp(-^t \mu x) \, dx,$$
where $dx$ is the uniform measure on $S^{n-1}$, which is $G$-invariant. Since $S^{n-1}$ is compact, we have $\Theta = \{ \theta = \mu \in \mathbb{R}^n \}$. Therefore we obtain the family of von Mises–Fisher distributions on $S^{n-1}$ from Fact 4.14.

**Fact 4.14** [4, Example 8.3] For $\mu \in \mathbb{R}^n$,

$$
\int_{S^{n-1}} \exp(-t \mu x) dx = (2\pi)^{\frac{n}{2}} \| \mu \|^{1-\frac{n}{2}} I_{\frac{1}{2}-1}(\| \mu \|).
$$

Here $dx$ is the SO$(n)$-invariant measure on $S^{n-1}$ with $\int_{S^{n-1}} dx = \frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}$ and $I_m(r)$ denotes the modified Bessel function of the first kind of order $m$ (see (10)).

### 4.7 Fisher–Bingham distribution

In this subsection, we realize the family of Fisher–Bingham distributions (see [11] for example) on the sphere $S^{n-1}$ by using a representation of SO$(n)$ for a positive integer $n \geq 2$.

**Proposition 4.15** Put $G = SO(n)$ and $H = SO(n-1) = \left\{ \begin{pmatrix} 1 & k \\ 0 & I_{n-1} \end{pmatrix} \mid k \in SO(n-1) \right\} \subset G$. Let $\rho: G \to GL(\mathbb{R}^n \oplus \text{Sym}(n, \mathbb{R}))$ be a representation given as follows:

$$
\rho(g)(v, S) := (gv, gS^t g) \quad ((v, S) \in \mathbb{R}^n \oplus \text{Sym}(n, \mathbb{R})).
$$

Then $v_0 := (e_1, E_{11}) \in \mathbb{R}^n \oplus \text{Sym}(n, \mathbb{R})$ is $H$-fixed, and by our method we obtain the family of Fisher–Bingham distributions on $S^{n-1}$,

$$
\{ c_{\mu, A} \exp(-t \mu x - tAx) dx \}_{(\mu, A) \in \mathbb{R}^n \oplus \text{Sym}(n, \mathbb{R})},
$$

where $c_{\mu, A}$ is the corresponding normalizing constant.

**Remark 4.16** We obtain the family of Kent distributions on $S^{n-1}$ as a subfamily of (12) if we assume the constraint $Ax = 0$.

**Proof** We use the identification $SO(n)/SO(n-1) \simeq S^{n-1}$, $gSO(n-1) \mapsto ge_1$. Since $G = SO(n)$ is compact, from Fact 4.1(ii), we obtain $W_0(G, H) = \{0\}$. We identify $(\mathbb{R}^n \oplus \text{Sym}(n, \mathbb{R}))^\vee$ with $\mathbb{R}^n \oplus \text{Sym}(n, \mathbb{R})$ by taking the direct sum of the standard inner product on $\mathbb{R}^n$ and the following inner product on $\text{Sym}(n, \mathbb{R})$:

$$
\langle x, y \rangle = \text{Tr}(xy) \quad (x, y \in \text{Sym}(n, \mathbb{R})).
$$

Then we have for $x = ge_1 \in S^{n-1} (g \in SO(n))$,

$$
d\tilde{p}_{(\mu, A)}(x) = \exp(-\langle (\mu, A), \rho(g)(e_1, E_{11}) \rangle) dx
\begin{align*}
&= \exp(-t \mu x - \text{Tr}(Ax^t x)) dx \\
&= \exp(-t \mu x - tAx) dx.
\end{align*}
$$
Since $S^{n-1} = G/H$ is compact, we have $\Theta = \mathbb{R}^n \oplus \text{Sym}(n, \mathbb{R})$. Therefore we obtain the family of Fisher–Bingham distributions \((12)\).

4.8 Hyperboloid distribution

In this subsection, we realize the family of hyperboloid distributions \([10]\) on the \(n\)-dimensional hyperbolic space \(H^n\) (see Definition 4.19) by using the natural representation of \(SO_0(1, n)\).

**Proposition 4.17** Let \(n\) be a positive integer and put \(G = SO_0(1, n)\) and \(H = SO(n)\). Let \(\rho : G \to GL(\mathbb{R}^{n+1})\) be the natural representation. We denote by \(e_0, \ldots, e_n\) the standard basis of \(\mathbb{R}^{n+1}\). Then \(v_0 := e_0 \in \mathbb{R}^{n+1}\) is \(H\)-fixed, and by our method we obtain the family of hyperboloid distributions on the \(n\)-dimensional hyperbolic space \(H^n\),

\[
\left\{ c_\kappa \exp(\kappa \langle \xi, v \rangle) d\nu(v) \right\}_{(\kappa, \xi) \in \mathbb{R}^+ \times H^n} (v \in H^n),
\]

where \(\langle \cdot, \cdot \rangle\) denotes a pseudo inner product given as \((15)\) below, \(\nu\) is a \(G\)-invariant measure on \(H^n\), which is described as \(\sinh^{n-1} u \sin^{n-2} v_1 \cdots \sin v_{n-2} du dv_1 \cdots dv_{n-1}\) in hyperbolic-spherical coordinates \((u, v_1, \ldots v_{n-1})\) (see section 2C in \([10]\) for details), and \(c_\kappa\) is the normalizing constant written as follows:

\[
c_\kappa = \frac{\kappa^{n-1}}{(2\pi)^{\frac{n-1}{2}} 2K_{\frac{n-1}{2}}(\kappa)}.
\]

Here \(K_\lambda\) is the modified Bessel function of the second kind and with index \(\lambda\).

**Remark 4.18** (See Section 7 in \([3]\), see also \([10]\)) In \([10]\), he denotes by \(H^n\) the \((n-1)\)-dimensional hyperbolic space.

- In the case \(n = 1\), we have \(c_\kappa = \frac{1}{2K_0(\kappa)}\).
- In the case \(n = 2\), we have \(c_\kappa = \frac{\kappa e^\kappa}{2\pi}\).

**Definition 4.19** (Hyperbolic space) Let \(n\) be a positive integer. The following manifold \(H^n\) with the metric which is the restriction to \(H^n\) of \(-dx_0^2 + dx_1^2 + \cdots + dx_n^2\) on \(\mathbb{R}^{n+1}\) is called \(n\)-dimensional hyperbolic space.

\[
H^n = \{ x \in \mathbb{R}^{n+1} \mid x_0 > 0, \langle x, x \rangle = -1 \}.
\]

Here \(\langle \cdot, \cdot \rangle\) is a pseudo inner product on \(\mathbb{R}^{n+1}\) given as follows:

\[
\langle x, y \rangle := -x_0y_0 + x_1y_1 + \cdots + x_ny_n.
\]

A Lie group \(SO_0(1, n)\), which is the identity component of \(SO(1, n)\), acts transitively on the \(n\)-dimensional hyperbolic space \(H^n\). Since the stabilizer of \(e_1 \in H^n\) is...
\[ SO(n), \text{ we identify } H^n \text{ with the homogeneous space } SO_0(1, n)/SO(n) \text{ as follows:} \]
\[ SO_0(1, n)/SO(n) \to H^n \]
\[ gSO(n) \mapsto g \cdot e_0. \]

**Proof of Proposition 4.17** We identify \((\mathbb{R}^{n+1})^\vee\) with the pseudo inner product (15). We have \(W_0(G, H) = \{0\}\) from Fact 4.1(iii). Take the \(G\)-invariant measure \(\nu\) on \(H^n\) as in Proposition 4.17. Then we have
\[ d\tilde{p}_y(v) = \exp(-\langle y, v \rangle) d\nu(v) \quad (y \in \mathbb{R}^{n+1}, v = ge_0 \in H^n). \]

**Fact 4.20** \([3, 10]\) \(\Theta = \{y \in \mathbb{R}^{n+1} | y_0 < 0, \langle y, y \rangle < 0\}\).

Since there is a natural bijection \(\mathbb{R}_{>0} \times H^n \to \Theta, (\kappa, \xi) \mapsto -\kappa \xi\), we use \(\mathbb{R}_{>0} \times H^n\) as a parameter space.

**Fact 4.21** (Section 7 in [3]) The normalizing constant of \(\exp(-\langle y, v \rangle) d\nu(v) = \exp(\kappa \langle \xi, v \rangle) d\nu(v)\) is given as (14).

Therefore, we obtain the family of hyperboloid distributions (13).

**4.8.1 A family of distributions on the upper half plane**

In this subsubsection, via a diffeomorphism between \(H^2\) and the upper half plane \(H\), we realize the corresponding family of distributions on the upper half plane to the family above of hyperboloid distributions for \(n = 2\). Since the exponential family on the upper half plane is compatible with the Poincaré metric, in this paper, we call it the family of **Poincaré distributions**, which is given as follows:
\[ \left\{ \frac{D e^{2D}}{\pi} \exp \left(-\frac{a(x^2 + y^2) + 2bx + c}{y^2} \right) \frac{dxdy}{y^2} \right\}_{(a \ b \ c) \in \text{Sym}^+(2, \mathbb{R})}, \]

where \(D = \sqrt{ac - b^2}\).

First, let us recall the definition of the upper half plane.

**Definition 4.22** (Upper half plane) The following manifold \(H\) with Poincaré metric \(\frac{dx^2 + dy^2}{y^2}\) is called upper half plane:
\[ H = \{x + iy \in \mathbb{C} | y > 0\}. \]

The Lie group \(SL(2, \mathbb{R})\) acts transitively on the upper half plane \(H\) by the Möbius transformations given by:
\[ g \cdot z := \frac{az + b}{cz + d} \quad \left( g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{R}) \right). \]
Since the stabilizer of $i$ is $SO(2)$, we identify $\mathcal{H}$ with the homogeneous space $\text{SL}(2, \mathbb{R})/SO(2)$ as follows:

$$\mathcal{H} \leftrightarrow \text{SL}(2, \mathbb{R})/SO(2), \quad (19)$$

$$z = x + iy \mapsto \begin{pmatrix} \sqrt{y} & x \sqrt{y} \\ 0 & \frac{1}{\sqrt{y}} \end{pmatrix} \text{SO}(2), \quad (20)$$

$$g \cdot i \leftrightarrow g \text{SO}(2). \quad (21)$$

Moreover, we use the following diffeomorphism:

$$\text{SL}(2, \mathbb{R})/SO(2) \rightarrow SO_0(1, 2)/SO(2), \quad (22)$$

which is induced by the following 2 : 1 group homomorphism:

$$\text{SL}(2, \mathbb{R}) \rightarrow SO_0(1, 2), \quad g \mapsto \text{Ad}(g). \quad (23)$$

Here since $\text{SL}(2, \mathbb{R})$ is connected and the adjoint representation $\text{Ad}: \text{SL}(2, \mathbb{R}) \rightarrow GL(\text{sl}(2, \mathbb{R}))$ preserves the pseudo inner product $\frac{1}{2} \text{Tr}(xy)$ with signature $(2, 1)$ on $\text{sl}(2, \mathbb{R})$, the above map is well-defined. We use a basis $f_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $f_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $f_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ of $\text{sl}(2, \mathbb{R})$. Since $SO(2) \subset \text{SL}(2, \mathbb{R})$ stabilizes $f_1$, we have $\text{Ad}(SO(2)) = \left\{ \begin{pmatrix} 1 \\ g \end{pmatrix} \in SO_0(1, 2) \mid g \in SO(2) \right\}$. From the composition of (19), (22) and (16), we have the following correspondence:

$$\mathcal{H} \rightarrow H^2, \quad (24)$$

$$x + iy \mapsto \frac{1}{2y} \begin{pmatrix} 1 + x^2 + y^2 \\ 1 - (x^2 + y^2) \end{pmatrix} = \begin{pmatrix} v_0 \\ v_1 \\ v_2 \end{pmatrix}. \quad (25)$$

Therefore, by putting $a = \frac{\xi_0 + \xi_1}{2\pi} \kappa$, $b = \frac{\xi_2}{2} \kappa$ and $c = \frac{\xi_0 - \xi_1}{2} \kappa$, we obtain the Poincaré distribution (18) from the hyperboloid distribution $\left\{ \frac{\kappa e^\kappa}{2\pi} \exp(\kappa(-\xi_0 v_0 + \xi_1 v_1 + \xi_2 v_2)) d\nu(v) \right\}_{(\kappa, \xi) \in \mathbb{R} > 0 \times H^2}$.

Acknowledgements The authors would like to thank Professors Hideyuki Ishi, Frédéric Barbaresco, Toshiyuki Kobayashi, Kenichi Bannai, Kei Kobayashi and Shintaro Hashimoto for valuable comments. The authors would also like to thank Atsushi Suzuki, Kenta Oono and Kentaro Minami for helpful comments and discussion. This work was supported by JST, ACT-X Grant Number JPMJAX190K, Japan.
Data Availability  Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Declarations

Conflict of interest  The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Barndorff-Nielsen, O.E.: Exponential Families-Exact Theory. Various Publication Series, vol. 19. Aarhus Universitet, Aarhus (1970)
2. Barndorff-Nielsen, O.E.: Information and Exponential Families in Statistical Theory. Wiley, Chichester (1978)
3. Barndorff-Nielsen, O.E.: Hyperbolic distributions and distribution on hyperbolae. Scand. J. Stat. 8, 151–157 (1978)
4. Barndorff-Nielsen, O.E., Blaesild, P., Eriksen, P.S.: Decomposition and Invariance of Measures, and Statistical Transformation Models. Springer, New York (1989)
5. Barndorff-Nielsen, O.E., Blaesild, P., Jensen, J.L., Jorgensen, B.: Exponential transformation models. Proc. R. Soc. Lond. A 379, 41–65 (1982)
6. Cohen, T.S., Welling, M.: Harmonic exponential families on manifolds. In: Proceedings of the 32nd International Conference on Machine Learning (ICML), vol. 37, pp. 1757–1765 (2015)
7. Diaconis, P., Ylvisaker, F.: Conjugate priors for exponential families. Ann. Stat. 7(2), 269–281 (1979)
8. Ingham, A.E.: An integral which occurs in statistics. Proc. Camb. Philos. Soc. 29, 271–276 (1933)
9. Ishi, H.: Homogeneous cones and their application to statistics. In: Graczyk, P., Hassairi, A. (eds.) Modern Methods of Multivariate Statistics. Lecture Notes of the CIMPA-FECYT-UNESCO-ANR Workshop held in Hammamet, September 2011. Travaux en Cours, vol. 82. Hermann Editeurs, Paris (2014)
10. Jensen, J.L.: On the hyperboloid distribution. Scand. J. Stat. 8, 193–206 (1981)
11. Kume, K., Walker, S.G.: On the Fisher–Bingham distribution. Stat. Comput. 19, 167–172 (2009)
12. Siegel, G.L.: Üer die analytische tehorie der quadratische Formen. Ann. Math. 36, 527–606 (1935)
13. Tojo, K., Yoshino, T.: Harmonic exponential families on homogeneous spaces. Inf. Geom. 4, 215–243 (2021)
14. Wishart, J.: The generalized product moment distribution in samples from a normal multivariate population. Biometrika 20A, 32–52 (1928)

Publisher’s Note  Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.