On the Positive von Mises–Fisher Distribution

Faculty Development Leave Presentation

Jose Guardiola

Department of Mathematics and Statistics
Texas A&M University–Corpus Christi

September 16, 2026

Who I Am

- **Jose Guardiola**; I joined Texas A&M University–Corpus Christi in 2004. As a professor of statistics, I have taught courses from introductory statistics to advanced graduate levels.
- **My research** focuses on distribution theory and probability models for constrained spaces, including hyperspheres, positive orthants, and simplexes, with applications in directional statistics, compositional data, and the medical, biological, and marine sciences.
- This **Faculty Development Leave** supported joint work with Dr. Eduardo García-Portugués at the Department of Statistics, Universidad Carlos III de Madrid, Leganés Campus.

Jose Guardiola

Motivation

Many modern data sets are naturally nonnegative and constrained:

$$\boldsymbol{x} \in \mathbb{S}_{+}^{p-1} = \left\{ \boldsymbol{x} \in \mathbb{R}_{+}^{p} : \|\boldsymbol{x}\| = 1 \right\}.$$

Examples include:

- compositional data after square-root transformation,
- microbiome relative abundances,
- sparse text vectors and images,
- environmental and marine science proportions.

Challenges for existing probability distributions:

- Waste probability mass on infeasible regions,
- The exact zeros are intractable or ignored.

Existing Distributions

Traditional von Mises–Fisher distribution (vMF)Mardia & Jupp (2000):

$$f_{\mathrm{vMF}}(\mathbf{x};\boldsymbol{\mu},\kappa) = c_{\mathrm{vMF},p}(\kappa)\exp\{\kappa\boldsymbol{\mu}^\top\mathbf{x}\}, \qquad c_{\mathrm{vMF},p}(\kappa) = \frac{\kappa^{p/2-1}}{(2\pi)^{p/2}\mathcal{I}_{p/2-1}(\kappa)},$$

$\boldsymbol{\mu}$: mean direction; κ: concentration. Limitation: Support is the full sphere, not only $\mathbb{S}_+^{p-1}$.

Spherical–Dirichlet distribution (SDD) (Guardiola, 2020, 2025):

$$f_{\mathrm{SDD}}(\mathbf{x};\boldsymbol{\alpha}) = 2^{p-1}\Gamma(\alpha_0)\prod_{i=1}^{p}\frac{x_i^{2\alpha_i-1}}{\Gamma(\alpha_i)}, \qquad \alpha_0 = \sum_{i=1}^{p}\alpha_i,$$

for $\mathbf{x}\in\mathbb{S}_+^{p-1}$, $\alpha_i > 0$. Limitation: Structural zeros require adjustment.

Foundational Work: The Positive vMF Distribution

The initial form of this distribution was developed in Jacob Harris's M.Sc. thesis at TAMUCC Harris (2024), which I chaired with Dr. Eduardo García-Portugués as co-chair.
Folding the ordinary vMF distribution onto the positive orthant gives

$$f_{\mathrm{vMF}_+}(\boldsymbol{x}; \boldsymbol{\mu}, \kappa) = 2^p c_{\mathrm{vMF},p}(\kappa) \prod_{j=1}^{p} \cosh(\kappa \mu_j x_j), \qquad \boldsymbol{x} \in \mathbb{S}_+^{p-1}.$$

- The 2^p-term folding sum collapses to a product of cosh terms
- This is algebraically exceptional and preserves closed-form normalization
- Evaluation scales linearly with dimension p
- The initial work established the distribution, but left open the **statistical theory needed for inference and applications**.

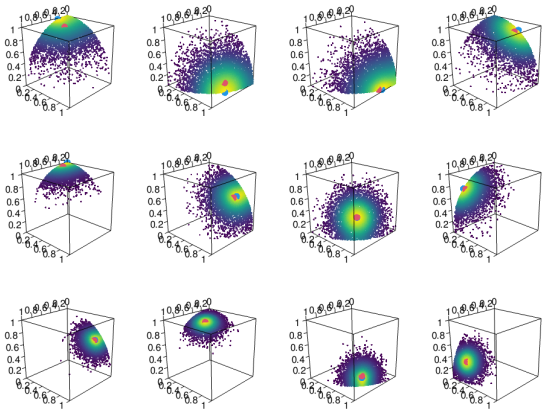
FDL Accomplishments: Statistical Theory

The FDL developed the initial distribution into a broader statistical framework:

1. Established identifiability and a natural parameterization
2. Developed likelihood inference and Fisher information
3. Established asymptotic inference for interior parameters
4. Characterized the mode and boundary activation
5. Derived explicit marginal densities, moments, and L^2 overlaps
6. Connected the model to simplex-valued compositional data
7. Developed a score test for uniformity on the simplex
8. Developed prediction regions, outlier tools, and density scoring
9. Demonstrated advantages for sparse, zero-rich data

Positive vMF on $\mathbb{S}^2_+$ ($p = 3$)

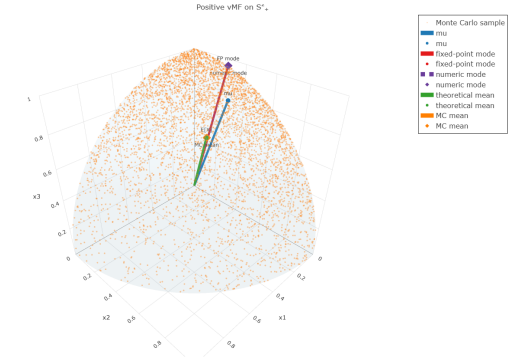
Examples of density shapes for different settings $(\boldsymbol{\mu}, \kappa)$ on the positive sphere



Varying $(\boldsymbol{\mu}, \kappa)$ produces a rich variety of modes and concentration patterns on $\mathbb{S}^2_+$.

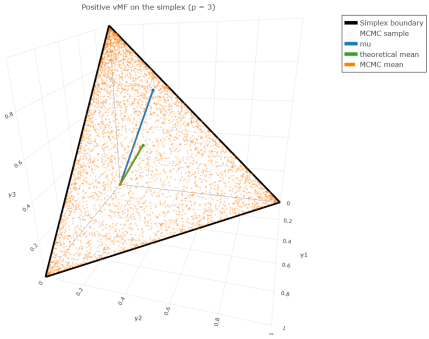
Positive vMF: Sphere and Simplex ($p = 3$)

On the Positive Sphere $\mathbb{S}^2_+$



- Samples from the positive vMF distribution.
- Mode and theoretical mean are displayed.
- Monte Carlo and theoretical quantities agree closely.

On the Simplex



- Squared coordinates induce a simplex model.
- The representation applies naturally to compositional data.
- Theoretical and Monte Carlo means agree closely.

FDL Accomplishments: Applications

We evaluated the methodology in settings where nonnegativity, sparsity, and boundary behavior are central:

- Glacial deposit compositions
- Microbiome compositions
- Sparse, high-dimensional document–term vectors

Applications included:

- Prediction regions and outlier detection
- Density-based model comparison and scoring
- Clustering and classification

FDL Accomplishments: Why This Can Change the Game

Classical compositional workflows often require special treatment of zeros. Here, zeros arise naturally as part of the geometry.

- Explicit, normalized probability density.
- Boundary observations, including exact zeros, are handled directly on $\mathbb{S}_+^{p-1}$.
- No artificial zero replacement is required.
- Folding preserves a closed-form normalized density.
- Likelihood-based inference remains available.
- The spherical model connects naturally to a simplex model for compositional data.

Core contribution

A tractable probability and inference framework for nonnegative directional and compositional data, with particular advantages for sparse, zero-rich data and density-based applications.

FDL Accomplishments: Research Outcomes

Journal Manuscripts

- Guardiola, J. H. & García-Portugués, E. (2026). *On the Positive von Mises–Fisher Distribution*. Manuscript submitted to the **top-tier statistics journal Biometrika**.
- Guardiola, J. H. & *Harris, J. (2026). *Comparing Directional Models for High-Dimensional Compositional Data: SDD and vMF Models*. Manuscript in preparation.

Conference Proceedings / Presentations

- Guardiola, J. H. & *Harris, J. (2026). *Comparing Directional Models for High-Dimensional Compositional Data: SDD and vMF Models*. **JSM 2026**, August 4, 2026, Boston MA.
- Guardiola, J. H. & García-Portugués, E. (2026). *On the Positive von Mises–Fisher Distribution*. **Math & Statistics Seminar**, April 3rd, 2026, TAMUCC.

FDL Accomplishments: Broader Research Impact

Research Collaboration

- Strengthened international collaboration with the Department of Statistics at **UC3M** and Dr. E. García-Portugués.
- Established a **research program** generating new thesis and dissertation opportunities in data science, mathematics, statistics, and marine science.
- Currently advising a PhD student in the CMSS program on a related applied research topic.

FDL Accomplishments: Future Research and Student Opportunities

The project now opens several research directions:

- **Mixture extensions:** clustering, classification, and density estimation for nonnegative directional data.
- **Bayesian extensions:** priors and posterior computation for $(\boldsymbol{\mu}, \kappa)$.
- **Boundary inference:** nonregular inference for inactive coordinates and structural zeros.
- **Applications:** environmental science, microbiome studies, text mining, and compositional data.
- **Student projects:** thesis and dissertation topics in CMSS, the DSc master's program, and the proposed PhD program in Data Science.

Long-term opportunity

This framework provides a platform for student-led research in modern constrained-data analysis.

Reflections from the Leave

- Focused research time made a challenging theoretical development possible.
- Face-to-face collaboration accelerated our research.
- The leave provided the continuity needed to develop an initial idea into a complete inferential framework.
- The collaboration strengthened professional relationships and opened new research opportunities.

Office at UC3M, Madrid

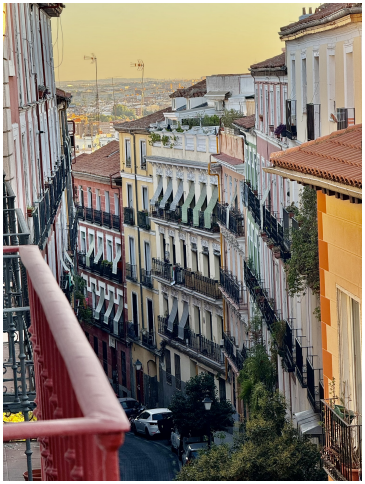
Closing Remarks: Thank You!

Takeaway

The Faculty Development Leave helped transform an initial idea into an inferential framework with practical applications and opportunities for future student research in mathematics, statistics, and data science.

Acknowledgments

This work was supported by a TAMU–CC Faculty Development Leave. I am grateful to the Department of Statistics at Universidad Carlos III de Madrid for its hospitality, to Eduardo García-Portugués for his collaboration, and to Jacob Harris, now a PhD student at Baylor University, for his early contributions.



Street view from Madrid apartment

References I

Guardiola, J. H. (2020). Spherical dirichlet distribution. *J. Stat. Distrib. Appl.* **7**, 1–13.

Guardiola, J. H. (2025). On the spherical dirichlet distribution: corrections and results. *arXiv preprint arXiv:2506.04441* .

Harris, J. (2024). *The Positive von Mises–Fisher Distribution*. Master's thesis, Texas A&M University–Corpus Christi.

Mardia, K. V. & Jupp, P. E. (2000). *Directional Statistics*. Wiley Series in Probability and Statistics. Chichester: John Wiley & Sons.