STA 437/2005: Methods for Multivariate Data

Week 5: Non-Gaussian Distributions

Piotr Zwiernik

University of Toronto

Table of contents

1. Elliptical distributions
   Spherical distributions
   Elliptical distributions

2. Copula models

Elliptical distributions

Why Study Elliptical Distributions?

- Generalize the multivariate normal distribution.
- Model data with heavy tails or outliers.
- Maintain symmetry and linear correlation structures.
- Applications in finance, insurance, and environmental studies.

Spherical Distributions

Orthogonal Matrices: $O(m) = \{U \in \mathbb{R}^{m\times m} : U^\top U = I_m\}$.

Spherical distribution

A random vector $X \in \mathbb{R}^m$ has a *spherical distribution* if for any $U \in O(m)$:

$$X \overset{d}{=} UX.$$

Characteristic function satisfies: $\psi_X(\boldsymbol{t}) = \psi_{UX}(\boldsymbol{t}) = \psi_X(U^\top \boldsymbol{t})$ and so **equivalently** $\psi_X(\boldsymbol{t})$ depends only on $\|\boldsymbol{t}\|$. Thus, the same applies to the density:

$$f_X(\mathbf{x}) \;=\; h(\|\boldsymbol{x}\|) \qquad \text{for some } h \text{ (generator).}$$

Examples of Spherical Distributions

Standard normal distribution $Z \sim N_m(0, I_m)$ is a simple example.

Spherical scale mixture of normals

If $Z \sim N_m(0, I_m)$ and a random variable $\tau > 0$ is independent of Z, then:

$$X = \frac{1}{\sqrt{\tau}} Z$$

has a spherical distribution.

Indeed: Let $U \in O(m)$, then

$$UX = \frac{1}{\sqrt{\tau}} UZ \overset{d}{=} \frac{1}{\sqrt{\tau}} Z = X.$$

Moment Structure of Spherical Distributions

Spherical symmetry implies:

$$\mathbb{E}[X] = 0,$$
$$\mathrm{var}(X) = cI_m, \quad \text{for some } c \geq 0.$$

Indeed: $\mathrm{var}(X) = \mathrm{var}(UX) = U\mathrm{var}(X)U^\top$ for any $U \in O(m)$

For $X = \frac{1}{\sqrt{\tau}}Z$ with $Z \sim N(0, I_m)$, $\tau > 0$, $\tau \perp\!\!\!\perp Z$:

$$\mathrm{var}(X) = \mathbb{E}[\tau^{-1}]I_m.$$

Indeed: $\mathbb{E}[X] = \mathbb{E}[\frac{1}{\sqrt{\tau}}Z] = \mathbb{E}[\frac{1}{\sqrt{\tau}}]\mathbb{E}[Z] = \mathbf{0}_m$ and so

$$\mathrm{var}(X) \;=\; \mathbb{E}XX^\top - \mathbb{E}[X]\mathbb{E}[X]^\top \;=\; \mathbb{E}[\tfrac{1}{\tau}ZZ^\top] \;=\; \mathbb{E}[\tfrac{1}{\tau}]\mathbb{E}[ZZ^\top] \;=\; \mathbb{E}[\tfrac{1}{\tau}]I_m$$

Independence of $\|X\|$ and $\frac{X}{\|X\|}$

Key Property

If X is spherical, the norm $\|X\| = \sqrt{X^\top X}$ is independent of the direction $\frac{X}{\|X\|}$.

Proof Sketch: Let $U \in O(m)$. Then:

$$\frac{X}{\|X\|} \stackrel{d}{=} \frac{UX}{\|UX\|} = U\frac{X}{\|X\|}.$$

The vector $\frac{X}{\|X\|}$ is rotationally invariant $\Longrightarrow$ has uniform distribution on the unit sphere (independent of what $\|X\|$ is).

A formal proof uses polar coordinates, see the notes.

Elliptical Distribution $E(\mu, \Sigma)$

Recall that $Z \sim N_m(\mathbf{0}_m, I_m)$ then $X = \mu + \Sigma^{1/2} Z \sim N_m(\mu, \Sigma)$.

Elliptical distribution

A random vector $X \in \mathbb{R}^m$ has an elliptical distribution if:

$$X = \mu + \Sigma^{1/2} Z,$$

where Z is a spherical random vector.

The density of $X \sim E(\mu, \Sigma)$ is of the form

$$f_X(\mathbf{x}) \;=\; c_m \sqrt{\det \Sigma^{-1}}\, h((\mathbf{x} - \mu)^\top \Sigma^{-1} (\mathbf{x} - \mu)).$$

The generator g controls the shape of the distribution (and its tails in particular).

Again: Why Elliptical Distributions?

- Generalize the multivariate normal distribution.
- Model data with heavy tails or outliers.
- Maintain symmetry and linear correlation structures.
- Applications in finance, insurance, and environmental studies.

Scale Mixtures of Normals

Scale mixture of normals is a special class of elliptical distributions.

Stochastic representation:

$$X = \mu + \frac{1}{\sqrt{\tau}} \Sigma^{1/2} Z,$$

where $Z \sim N_m(0, I_m)$ and $\tau > 0$ is independent of Z.

Special Cases of Scale Mixture of Normals

- $\tau \equiv 1$: Multivariate normal.
- $\tau \sim \frac{1}{k}\chi_k^2$: Multivariate t-distribution with k degrees of freedom.
  - Smaller k means heavier tails. Gaussian is the limit $k \to \infty$.
- $\tau \sim \text{Exp}(1)$: Multivariate Laplace.

Covariance and Correlation in Elliptical Distributions

Σ is called the **scale matrix**. It is generally not equal to the covariance matrix.

$$\mathrm{Var}(X) = c\Sigma, \quad c > 0.$$

Correlation structure is still governed by Σ:

$$R_{ij} \;=\; \frac{c\Sigma_{ij}}{\sqrt{c\Sigma_{ii}c\Sigma_{jj}}} \;=\; \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\Sigma_{jj}}}.$$

Similarly, if $X \sim E(\mu, \Sigma)$ and $X = (X_A, X_B)$ then

$$\mathbb{E}(X_A | X_B = \mathbf{x}_B) \;=\; \mathbb{E}(X_A) - \Sigma_{A,B}\Sigma_{B,B}^{-1}(\mathbf{x}_B - \mu_B)$$

exactly as in the Gaussian case.

Copula models

Cumulative Distribution Function (CDF)

Let $X = (X_1, \ldots, X_m)$ be a random vector. Its **CDF** is:

$$F(x_1, \ldots, x_m) = \mathbb{P}(X_1 \le x_1, X_2 \le x_2, \ldots, X_m \le x_m).$$

Marginal CDF: $F_1(x_1) = \mathbb{P}(X_1 \le x_1) = \lim_{x_2 \to \infty} \cdots \lim_{x_m \to \infty} F(x_1, x_2, \ldots, x_m)$.
(similar for any other margin)

If f is the corresponding density, then:

$$f(x_1, \ldots, x_m) \;=\; \frac{\partial^m}{\partial x_1 \cdots \partial x_m} F(x_1, \ldots, x_m)$$

$$F(x_1, \ldots, x_m) \;=\; \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_m} f(y_1, \ldots, y_m) \mathrm{d}y_1 \cdots \mathrm{d}y_m.$$

If $U \sim U[0,1]$ then $F(u) = u \;\; u \in [0,1]$.

What is a Copula?

- A **copula** is a function that captures the **dependence structure** between random variables, separate from their marginal distributions.

Definition

A function $C : [0,1]^m \to [0,1]$ is a **copula** if it is a CDF with uniform marginals, that is, $C_1(u_1) = u_1, \dots, C_m(u_m) = u_m$, where C_i are the marginal CDF's.

For example, the copula $C(\boldsymbol{u}) = u_1 \cdots u_m$ corresponds to a m independent $U[0,1]$.

Why use copulas?

- To model non-Gaussian dependencies.
- To analyze dependence independently of marginal behaviors.

Sklar's Theorem

Theorem (Sklar, 1959)

Let $X = (X_1, \ldots, X_m)$ be a random vector with joint CDF F and marginals $F_1, \ldots, F_m$. There exists a unique copula C such that:

$$F(x_1, \ldots, x_m) = C(F_1(x_1), \ldots, F_m(x_m)).$$

Conversely, given marginals $F_1, \ldots, F_m$ and a copula C, the joint CDF F is defined by the same formula.

- C captures **dependence structure**.
- $F_1, \ldots, F_m$ capture marginal behaviors.

Understanding Sklar's Theorem

- When $m = 1$, $C(u) = u$, the identity function on $[0, 1]$.
- If X is continuous with CDF F, then $F(X) \sim U(0, 1)$.

  Proof: If X is continuous, F is strictly increasing on the support. Hence

$$\mathbb{P}(F(X) \leq u) = \mathbb{P}(X \leq F^{-1}(u)) = F(F^{-1}(u)) = u.$$

Copulas from Uniform Marginals

- Let $X = (X_1, \ldots, X_m)$ with CDF F.
- Define $U_i = F_i(X_i)$, where F_i are the marginal CDFs.
- The transformed variables $U = (U_1, \ldots, U_m)$ have uniform marginals.

$$\mathbb{P}(U_1 \leq u_1, \ldots, U_m \leq u_m) = C(u_1, \ldots, u_m).$$

- Also note that

$$\mathbb{P}(U_1 \leq u_1, \ldots, U_m \leq u_m) = F(F_1^{-1}(u_1), \ldots, F_m^{-1}(u_m))$$

  and so the copula can be computed explicitly.
- Sklar's theorem ensures C is unique for continuous distributions.

Simple Example of a Copula

- Joint CDF:

$$F_{X,Y}(x,y) = \begin{cases} 0 & x < 0 \text{ or } y < 0, \\ x^2y^2 & 0 \leq x, y \leq 1, \\ 1 & x > 1 \text{ and } y > 1, \\ \min(x^2, y^2) & \text{otherwise.} \end{cases}$$

- Marginal CDFs:

$$F_X(x) = x^2, \quad F_Y(y) = y^2 \quad \text{for } 0 \leq x, y \leq 1.$$

- Copula:

$$C(u,v) = uv \quad \text{if } u, v \leq 1.$$

Invariance under Monotone Transformations

- Let $Y_i = f_i(X_i)$, where f_i are strictly increasing transformations.
- The copula remains unchanged.
- Proof outline:
  - Marginals transform: $G_i(y_i) = F_i(f_i^{-1}(y_i))$.
  - Copula representation remains:

$$C(u_1, \ldots, u_m) = F(F_1^{-1}(u_1), \ldots, F_m^{-1}(u_m)).$$

Density of a Copula

- The PDF of a copula C is obtained by differentiating its CDF:

$$c(\mathbf{u}) = \frac{\partial^m C(\mathbf{u})}{\partial u_1 \cdots \partial u_m}.$$

- Using the chain rule:

$$c(\mathbf{u}) = \frac{f(\mathbf{x})}{\prod_{i=1}^{m} f_i(x_i)},$$

where f is the joint density and f_i are marginal densities.

Gaussian Copula

- Derived from the multivariate normal distribution.
- Simplify to standard normal marginals:

$$c(\mathbf{u}; \Sigma) = \det(\Sigma)^{-1/2} \exp\left(-\frac{1}{2}\mathbf{x}^\top(\Sigma^{-1} - I_m)\mathbf{x}\right),$$

where $x_i = \Phi^{-1}(u_i)$ and Φ is the standard normal CDF.

Applications of Copulas

- **Finance:** Modeling dependencies in asset returns.
- **Insurance:** Understanding risks in correlated claims.
- **Environmental Science:** Joint modeling of extreme events (e.g., floods).
- **Medical Statistics:** Modeling dependence in survival times.