ACM Winter School 2025

Lecture 2: Copulas and Model Risk

© Anand Deo (Indian Institute of Management, Bangalore)

Recap of Session 1: Input Modelling

(Gaussian Distributions)

Task: Recall that bootstrap merely requires access to samples.

- So far: we can only model uniform samples of data
- Broader question —> how to sample from the JOINT distribution of losses?

Basic Model: Multivariate Normal Distribution.

- ξ is said to be multivariate normal if $\xi = \mu + CN$ where N is a d-dimensional vector of independent Gaussians and $C \in \mathbb{R}^{d \times d}$

- It turns out that $E[\xi] = \mu$ and Cov$(\xi) = \Sigma = CC^\intercal$. Notation: $\xi \sim$ MVN(μ, Σ)

Simulation of a MVN(μ, Σ) random vector:

- Generate d univariate Gaussian random variables $N = (N_1, \ldots, N_d)$
- Cholesky decomposition: $\Sigma = UDU^\intercal$ where $U \rightarrow$ upper-triangular. Set $C = U\sqrt{D}$
- Return $\xi = \mu + CN$

Capturing Extremes: The Elliptical Distribution

(Allowing arbitrary scale)

Task: Recall that bootstrap merely requires access to samples.

- So far: we can only model uniform samples of data
- Broader question —> how to sample from the JOINT distribution of losses?

Approach 1: Impose a parametric structure.

- Assume that $\xi = \mu + \Lambda U R$, where $\mu \in \mathbb{R}^d$ and U —> uniform distribution on the unit d-sphere and $R \rightarrow$ univariate distribution, independent of U

- ξ —> Elliptically distributed. Notation: $\xi \sim \text{Ell}(\mu, \Lambda, R)$

Procedure for simulation:

i) Generate d samples from a Gaussian: $N = (N_1, \ldots, N_d)$ and set $U = N/\|N\|$
ii) Sample R independently of N (how?)
iii) Return $\xi = \mu + \Lambda U R$ —> sample from $\mathtt{Ell}(\mu, \Lambda, R)$

Capturing Extremes: The Elliptical Distribution

(Allowing arbitrary scale)

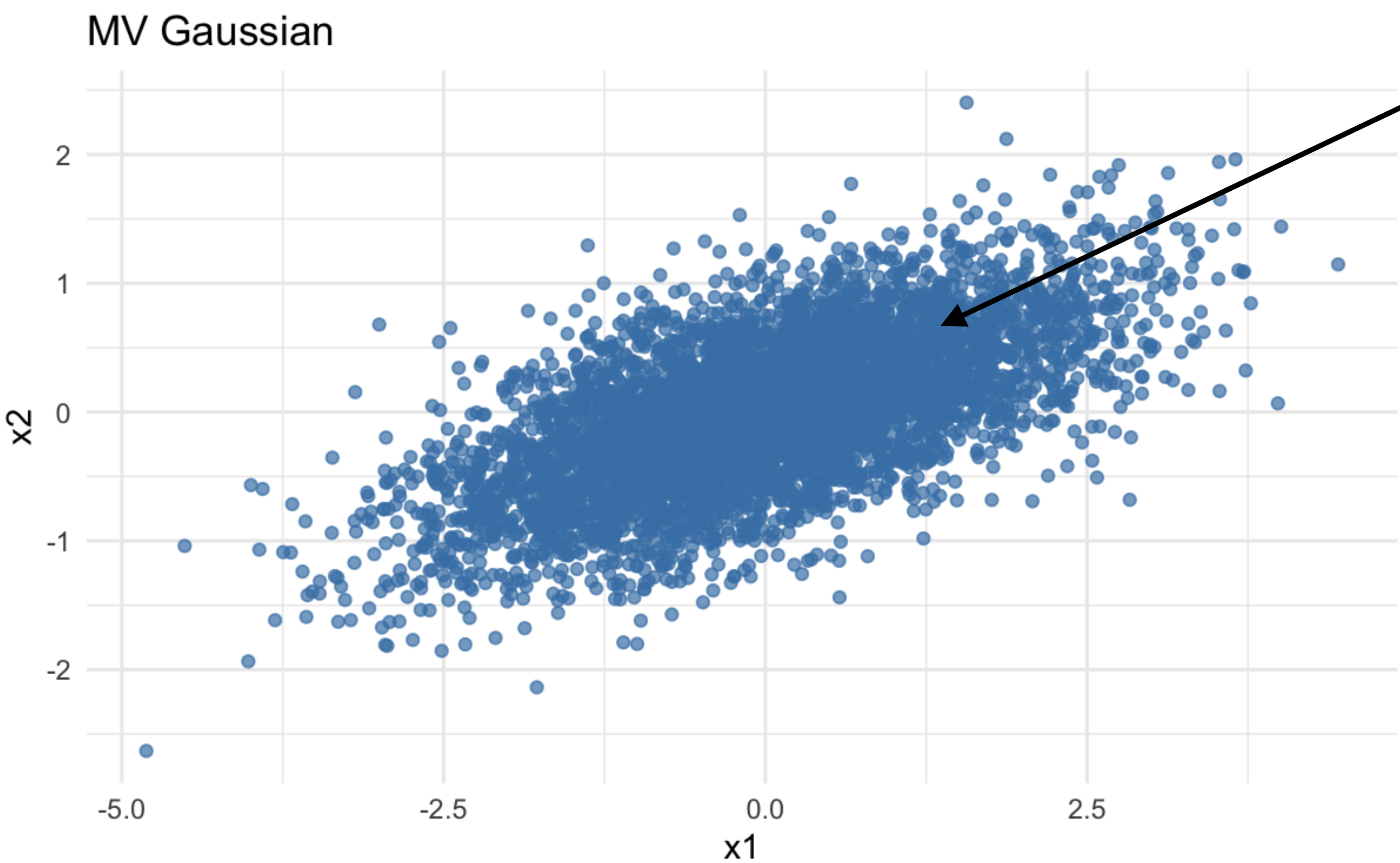
Task: Recall that bootstrap merely requires access to samples.

- So far: we can only model uniform samples of data
- Broader question —> how to sample from the JOINT distribution of losses?

Approach 1: Impose a parametric structure.

- Assume that $\xi = \mu + \Lambda U R$, where $\mu \in \mathbb{R}^d$ and U —> uniform distribution on the unit d-sphere and $R \rightarrow$ univariate distribution, independent of U

- ξ —> Elliptically distributed. Notation: $\xi \sim \text{Ell}(\mu, \Lambda, R)$



- Λ—> shape matrix (captures dependence)

- R—> tail behaviour (allows marginals other than Gaussian)

Capturing Extremes: The Elliptical Distribution

(Allowing arbitrary scale)

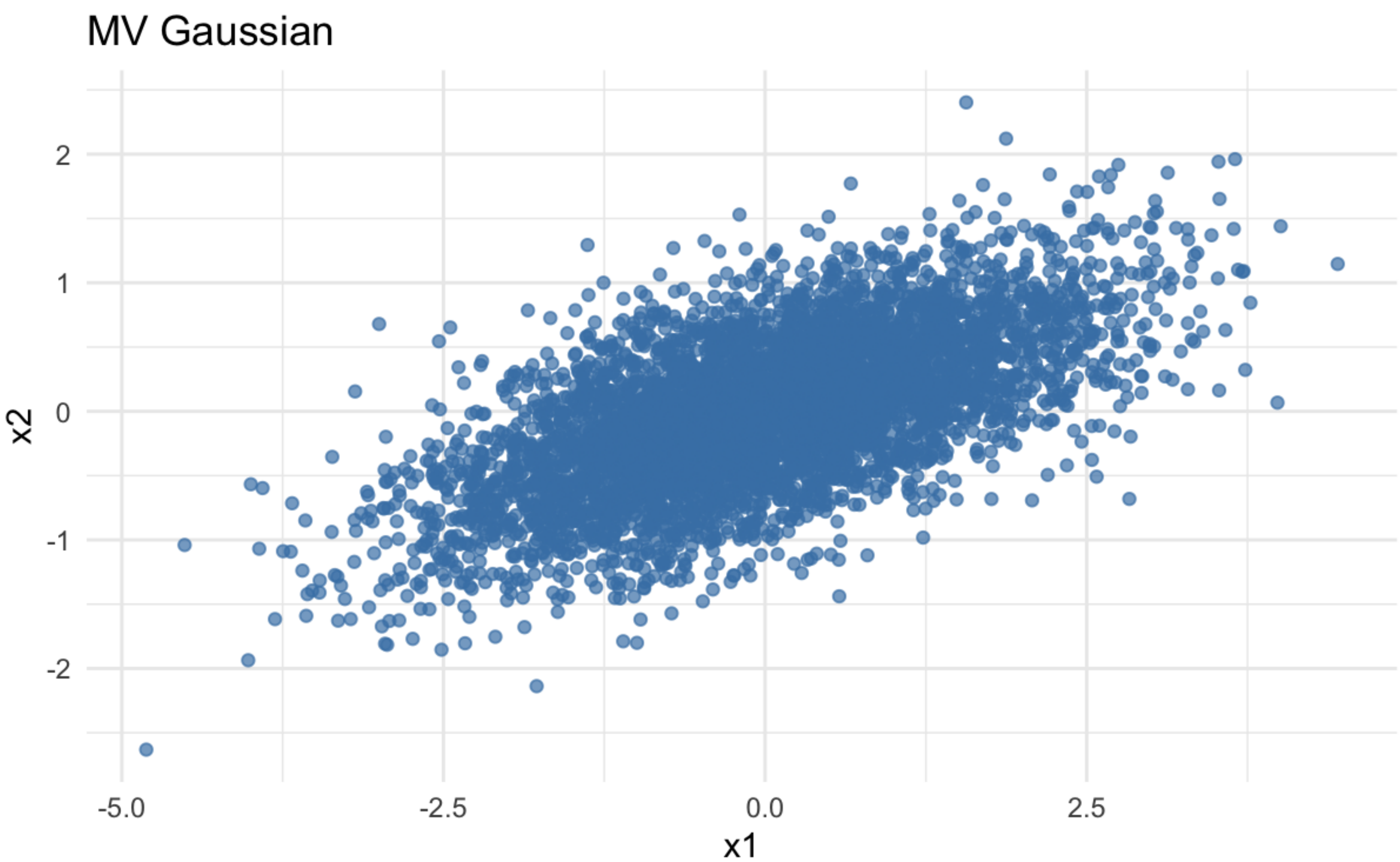
Task: Recall that bootstrap merely requires access to samples.

- So far: we can only model uniform samples of data
- Broader question —> how to sample from the JOINT distribution of losses?

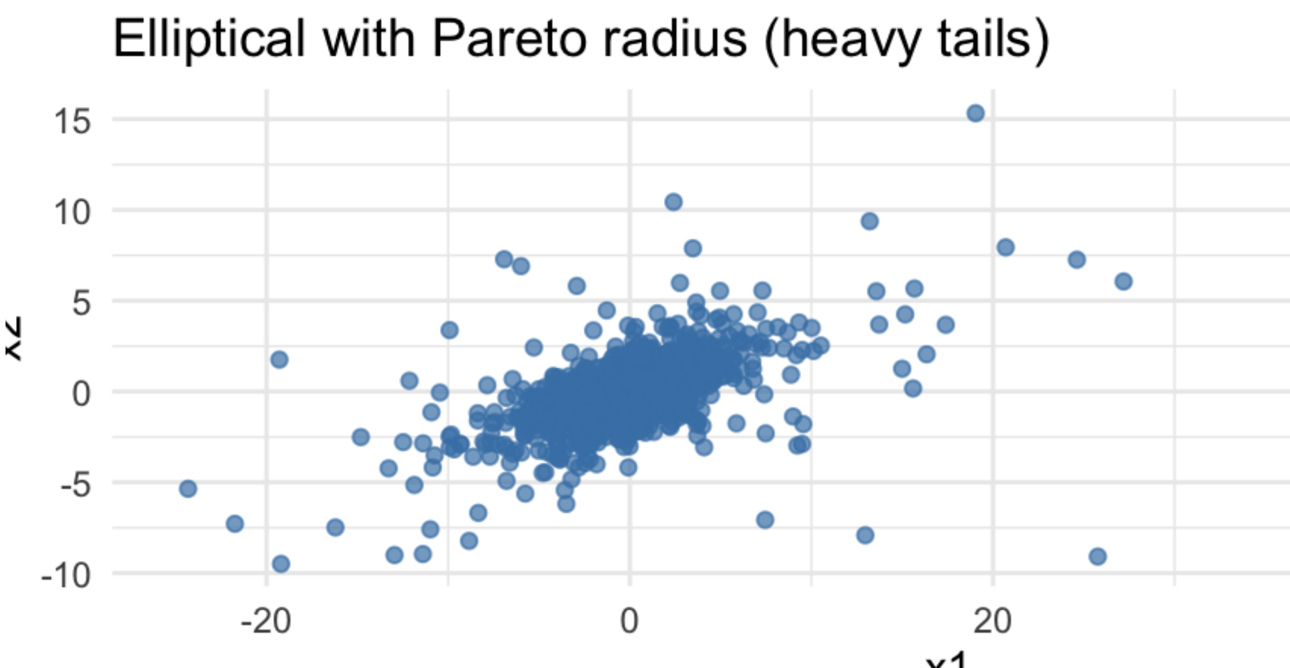
Approach 1: Impose a parametric structure.

- Assume that $\xi = \mu + \Lambda U R$, where $\mu \in \mathbb{R}^d$ and U —> uniform distribution on the unit d-sphere and $R \rightarrow$ univariate distribution, independent of U

- ξ —> Elliptically distributed. Notation: $\xi \sim \text{Ell}(\mu, \Lambda, R)$



R —> models tail co-movement



Capturing Extremes: The Elliptical Distribution

(Allowing arbitrary scale)

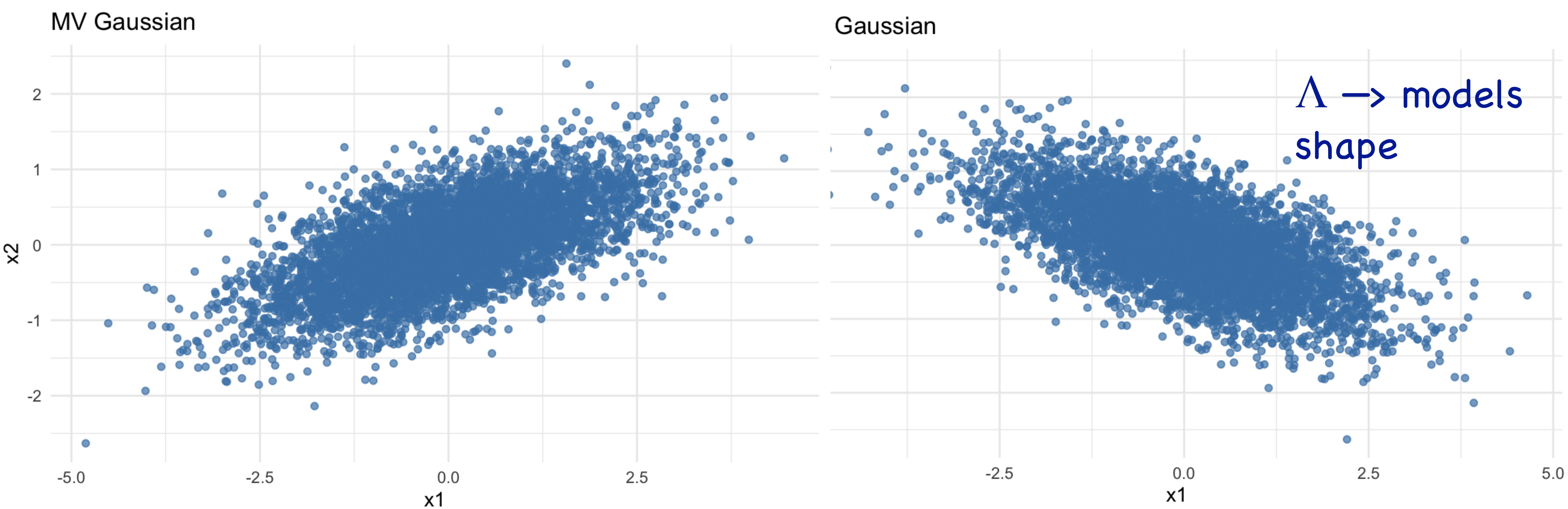
Task: Recall that bootstrap merely requires access to samples.

- So far: we can only model uniform samples of data
- Broader question —> how to sample from the JOINT distribution of losses?

Approach 1: Impose a parametric structure.

- Assume that $\xi = \mu + \Lambda U R$, where $\mu \in \mathbb{R}^d$ and U —> uniform distribution on the unit d-sphere and $R \rightarrow$ univariate distribution, independent of U

- ξ —> Elliptically distributed. Notation: $\xi \sim \text{Ell}(\mu, \Lambda, R)$



Capturing Extremes: The Elliptical Distribution

(Allowing arbitrary scale)

Task: Recall that bootstrap merely requires access to samples.

- So far: we can only model uniform samples of data
- Broader question –> how to sample from the JOINT distribution of losses?

Approach 1: Impose a parametric structure.

- Assume that $\xi = \mu + \Lambda U R$, where $\mu \in \mathbb{R}^d$ and U –> uniform distribution on the unit d-sphere and $R \to$ univariate distribution, independent of U
- ξ –> Elliptically distributed. Notation: $\xi \sim \text{Ell}(\mu, \Lambda, R)$

Limitations: Although simple in modelling terms there are several!

- All marginals are from the same family (Gaussian, Laplace, student-t...)
- Dependence is rigid and driven by Λ

These are due to the parametric nature of the model

Agenda for the course

Session 1

Input Modelling 🚂

- The bootstrap
- Inverse Transforms
- Sampling from multivariate distributions

Session 2

The Copula & Model Risk 🚂

- The copula: A non-parametric modelling tool

Copulas: Input Modelling, but for multivariate distributions

(Inverse Transform + Dependence = Multivariate Model)

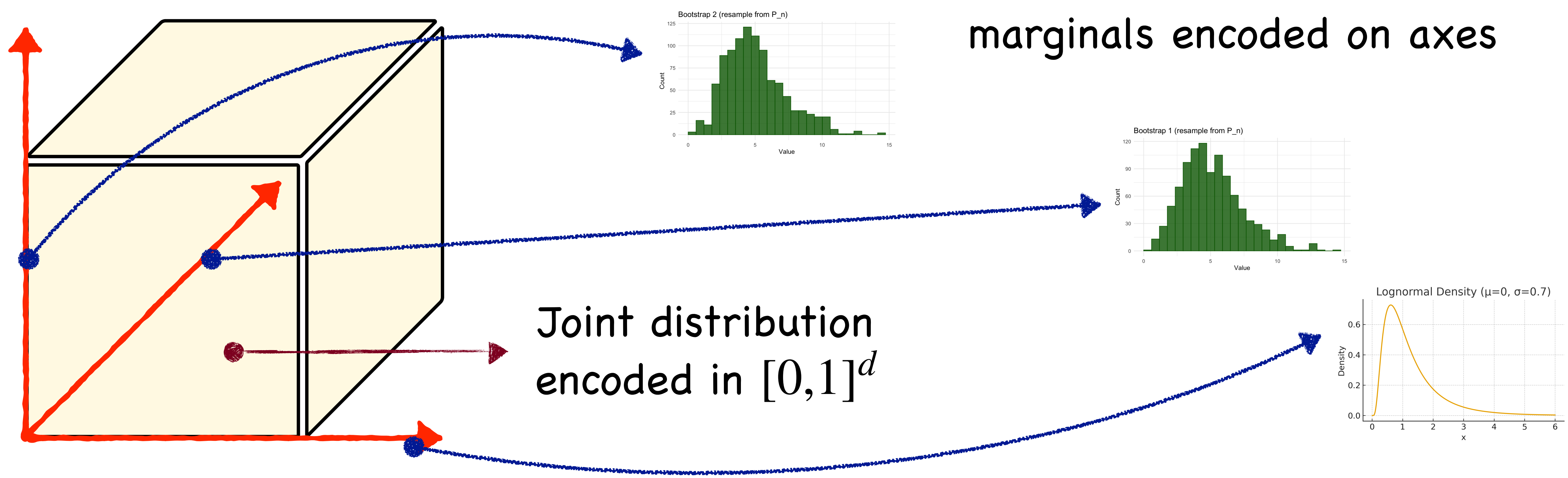
Approach 2: Copula → Random vector $\boldsymbol{U} \in [0,1]^d$ such that each U_i is Unif$(0,1)$

Copulas isolate the "dependance" in a random vector from its marginals

Definition: If $\xi \in \mathbb{R}^d$ be a random vector with continuous marginals $(F_1, \dots, F_d)$. The distribution of the random vector

$$\boldsymbol{U} = (U_1, \dots U_d) = (F_1(\xi_1), \dots, F_d(\xi_d))$$

is called the copula of ξ.



Copulas: Input Modelling, but for multivariate distributions

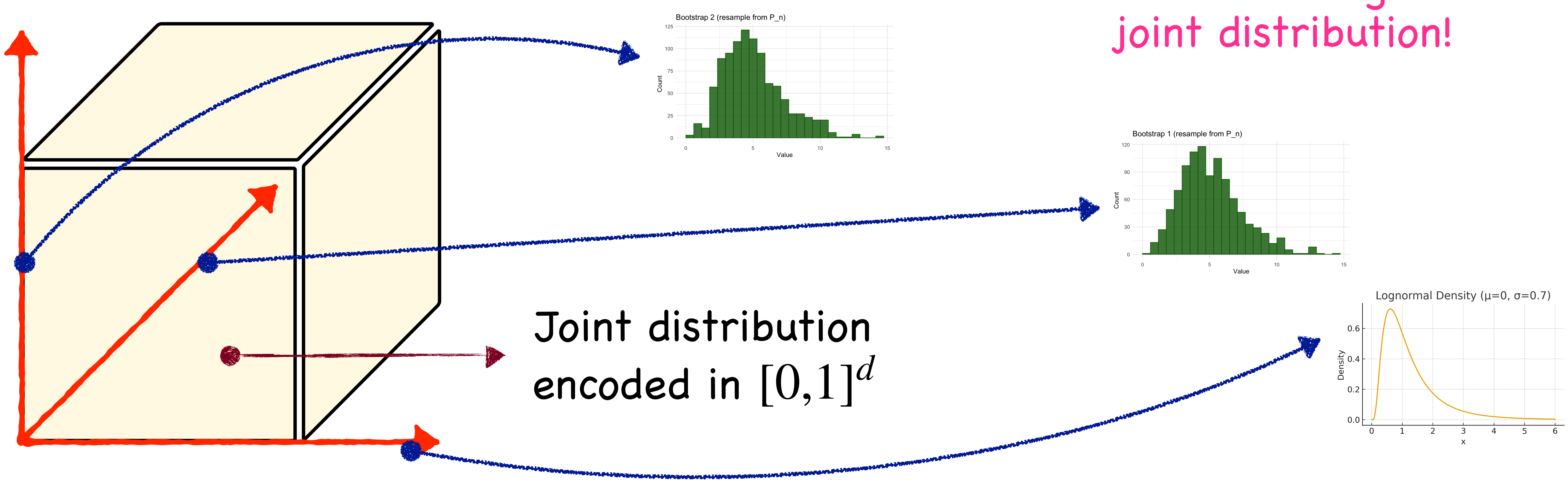
(Inverse Transform + Dependence = Multivariate Model)

Sklar's Theorem (invariance to marginal transformations): Let C be a copula. Let $\boldsymbol{U} \sim C$ and F_i be continuous cdfs. Set $\xi_i = F_i^{-1}(U_i)$. Then

$$\xi = (\xi_1, \ldots, \xi_d)$$

has marginal distribution $(F_1, \ldots, F_d)$ and the copula of ξ is $\boldsymbol{C}$.

Can capture any mix of continuous marginal + joint distribution!



Copulas: Input Modelling, but for multivariate distributions

(Inverse Transform + Dependence = Multivariate Model)

Illustration - Gaussian Copula: Let $\boldsymbol{Q} \sim \text{MVN}(0,\Sigma)$ such that $\Sigma_{ii} = 1$ for all i. Then a Gaussian (normal) copula can be generated using $\boldsymbol{Q}$ as follows:

$$\boldsymbol{U} = \left(\Phi(Q_1), \ldots \Phi(Q_d)\right) \text{ where } \Phi \rightarrow \text{ normal cdf}$$

In class example - Outline a procedure to sample from a bivariate Gaussian copula with a given covariance matrix, but whose marginals are F_1 and F_2

Solution:

- Generate a sample from $Q \sim \text{MVN}(0,\Sigma)$
- Set $U_i = \Phi(Q_i)$ for $i = 1,2$
- Output $\boldsymbol{R} = (F_1^{-1}(U_1), F_2^{-1}(U_2))$

Convince yourself that the sample generated has the desired marginal/joint distribution...

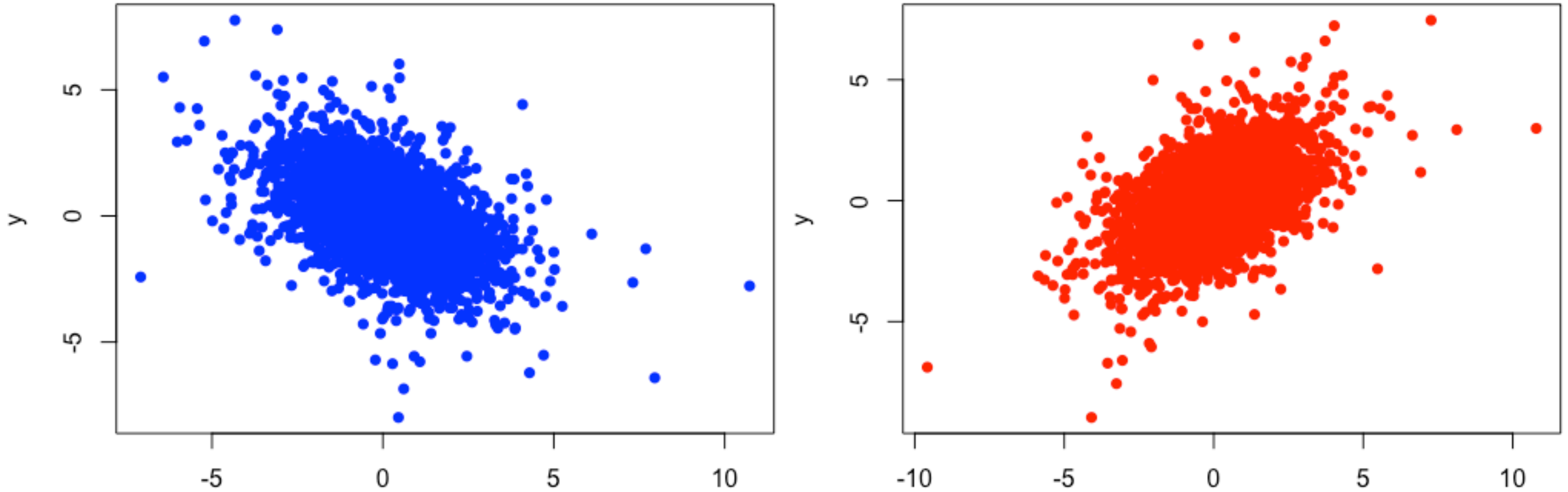
Copulas: Input Modelling, but for multivariate distributions

(Inverse Transform + Dependence = Multivariate Model)

Illustration - Gaussian Copula: Let $Q \sim \text{MVN}(0,\Sigma)$ such that $\Sigma_{ii} = 1$ for all i.
Then a Gaussian (normal) copula can be generated using Q as follows:

$$U = \left(\Phi(Q_1), \ldots \Phi(Q_d)\right) \text{ where } \Phi \rightarrow \text{ normal cdf}$$

Gaussian Copula with Laplace Marginals



Correlation = -0.8

Correlation = +0.8

Copulas: Input Modelling, but for multivariate distributions

(Inverse Transform + Dependence = Multivariate Model)

Illustration - Gaussian Copula: Let $\boldsymbol{Q} \sim \text{MVN}(0,\Sigma)$ such that $\Sigma_{ii} = 1$ for all i.
Then a Gaussian (normal) copula can be generated using $\boldsymbol{Q}$ as follows:

$$\boldsymbol{U} = \left(\Phi(Q_1), \ldots \Phi(Q_d)\right) \text{ where } \Phi \rightarrow \text{ normal cdf}$$

- Dependence is still captured parametrically (via Σ)
- This inherits the same limitations as elliptical models: e.g., no tail asymmetry, zero tail dependence, etc.

Next up: empirical / learned copulas – let the data specify the copula instead of imposing a Gaussian form

The empirical copula: modelling dependence from data

(Using DGMs to learn a non-parametric model)

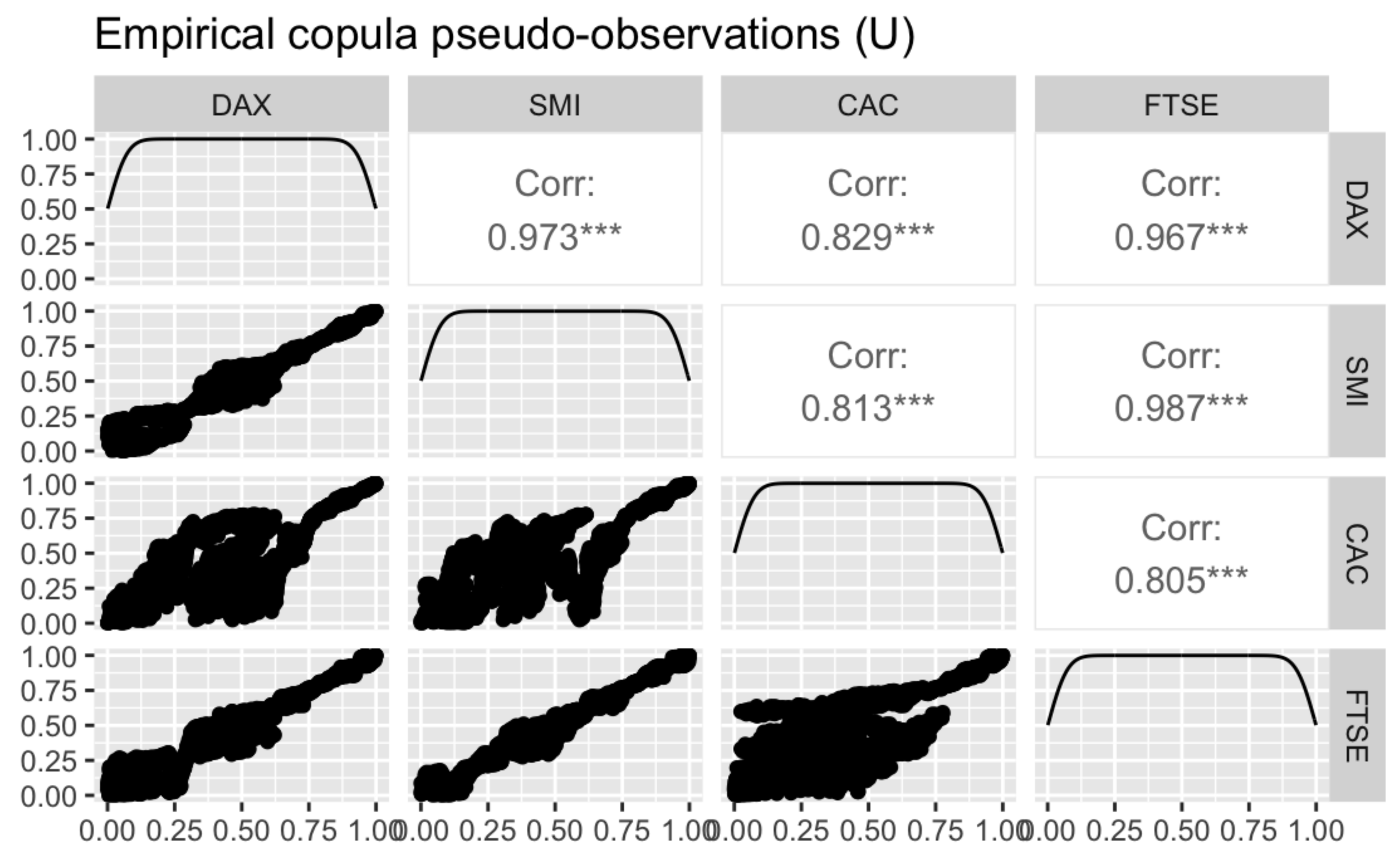
Definition: If $\xi \in \mathbb{R}^d$ be a random vector with continuous marginals $(F_1, \ldots, F_d)$. The distribution of the random vector

$$\boldsymbol{U} = (U_1, \ldots U_d) = (F_1(\xi_1), \ldots, F_d(\xi_d))$$

is called the copula of ξ.

Problem: $F_1, \ldots F_d$ unknown in practice

- Empirical cdf $F_{i,n}$ approximates F_i well for large n
- Do rank transform: $U_{i,k} = F_{i,n}(\xi_{i,k})$
- Empirical copula → Distribution of $(U_{1,k}, \ldots U_{d,k})$



Crucial: This is a fully non-parametric setup which only requires access to samples of data!

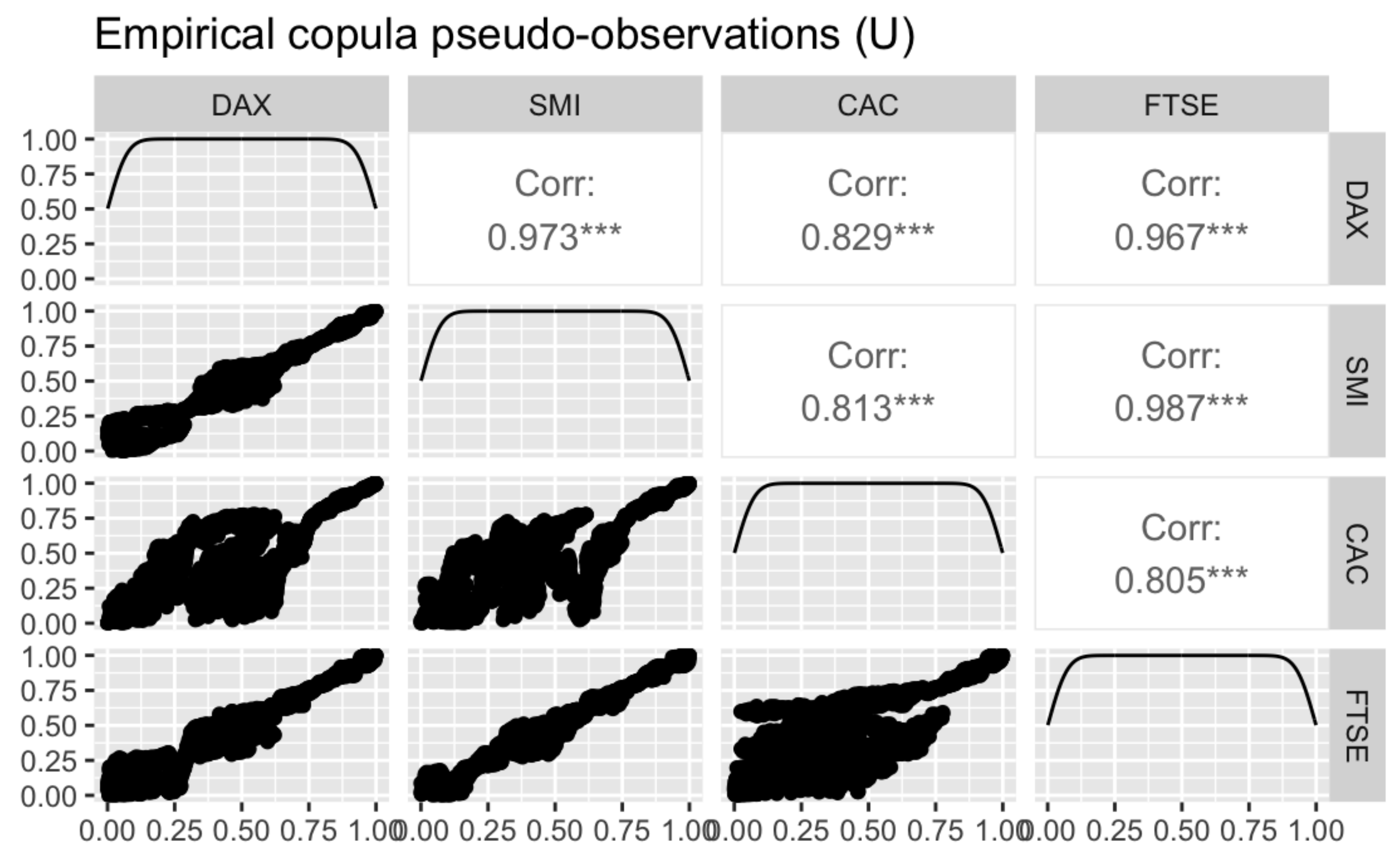
The empirical copula: modelling dependence from data

(Using DGMs to learn a non-parametric model)

Question: How does one learn the copula from data samples?

Problem: $F_1, \ldots F_d$ unknown in practice

- Empirical cdf $F_{i,n}$ approximates F_i well for large n
- Do rank transform: $U_{i,k} = F_{i,n}(\xi_{i,k})$
- Empirical copula $\rightarrow$ Distribution of $(U_{1,k}, \ldots U_{d,k})$

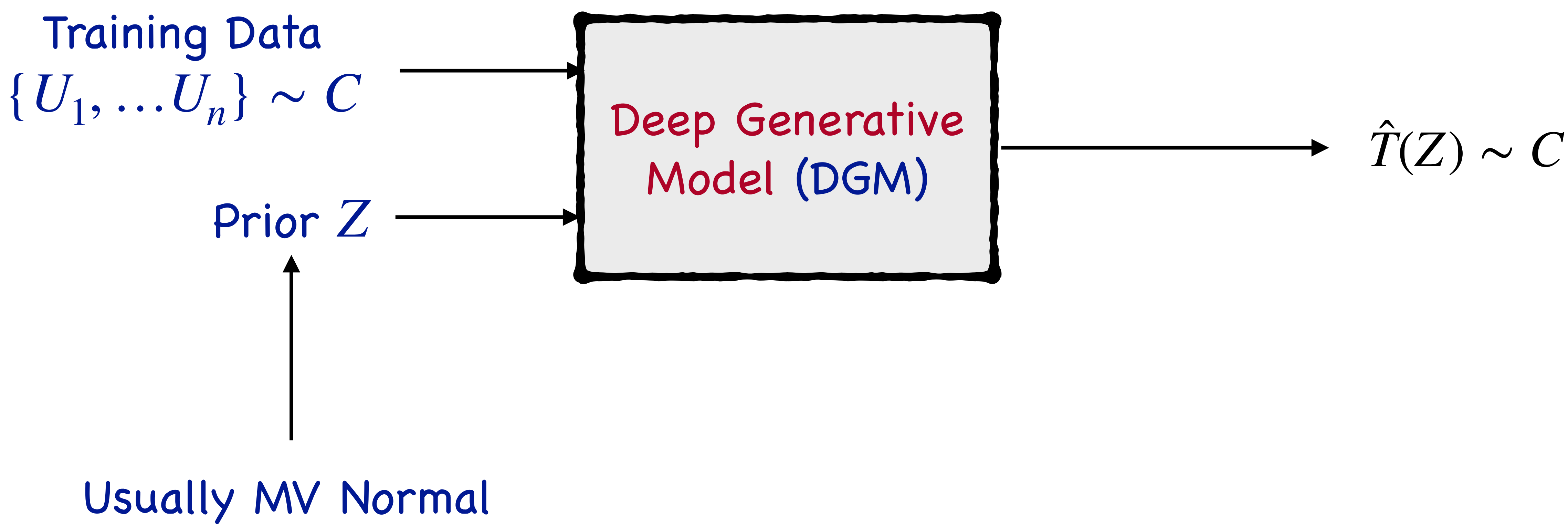


Crucial: This is a fully non-parametric setup which only requires access to samples of data!

Primer: Deep Generative Models

(Learning densities on compact sets)

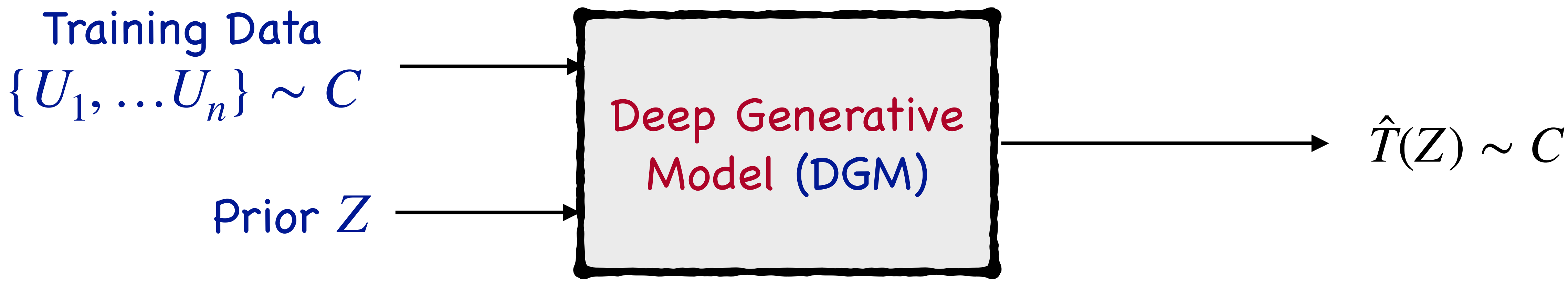
Problem: Suppose you are given samples $\{U_1, \ldots U_n\} \sim C$ with a density f_C. Can you use these samples to learn f_C?



Primer: Deep Generative Models

(Learning densities on compact sets)

Problem: Suppose you are given samples $\{U_1, \dots U_n\} \sim C$ with a density f_C. Can you use these samples to learn f_C?

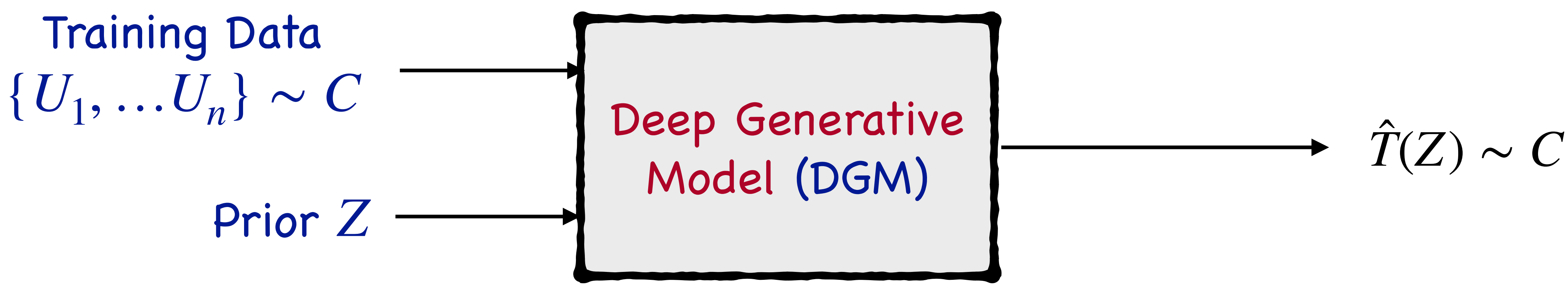


- DGMs —> large neural networks that learn a mapping $\hat{T}$ that take a (high-dimensional) prior Z to the target distribution C.
- Can approximate a rich class of non-parametric models! (See next slide for error bounds).
- In our context: use DGM to learn the copula.

Primer: Deep Generative Models

(Learning densities on compact sets)

Problem: Suppose you are given samples $\{U_1, \ldots U_n\} \sim C$ with a density f_C. Can you use these samples to learn f_C?



Theorem – Minimax learning rate (Liang, 2021): Let the target distribution have a density that is in the Sobolev class $W^\alpha(r)$. Then, for a suitable neural architecture, the following rates are achievable

$$\inf_{\nu_n} \sup_{\nu \in W^\alpha(r)} E[d_{TV}(\nu_n, \nu)] \asymp n^{-\frac{\alpha}{2\alpha + d}} \vee n^{-1/2}$$

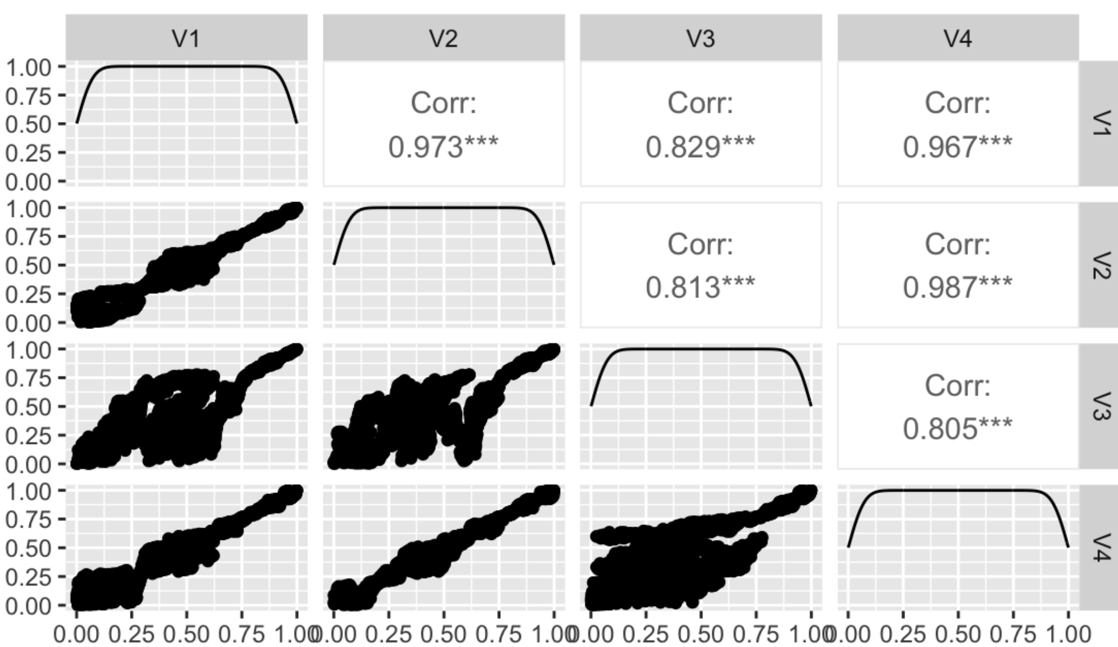
- Sobolev class α ≈ functions whose order α derivatives are square-integrable.
- Implication: If density is smooth, then DGMs can learn quite well!

The empirical copula: learning the copula using data

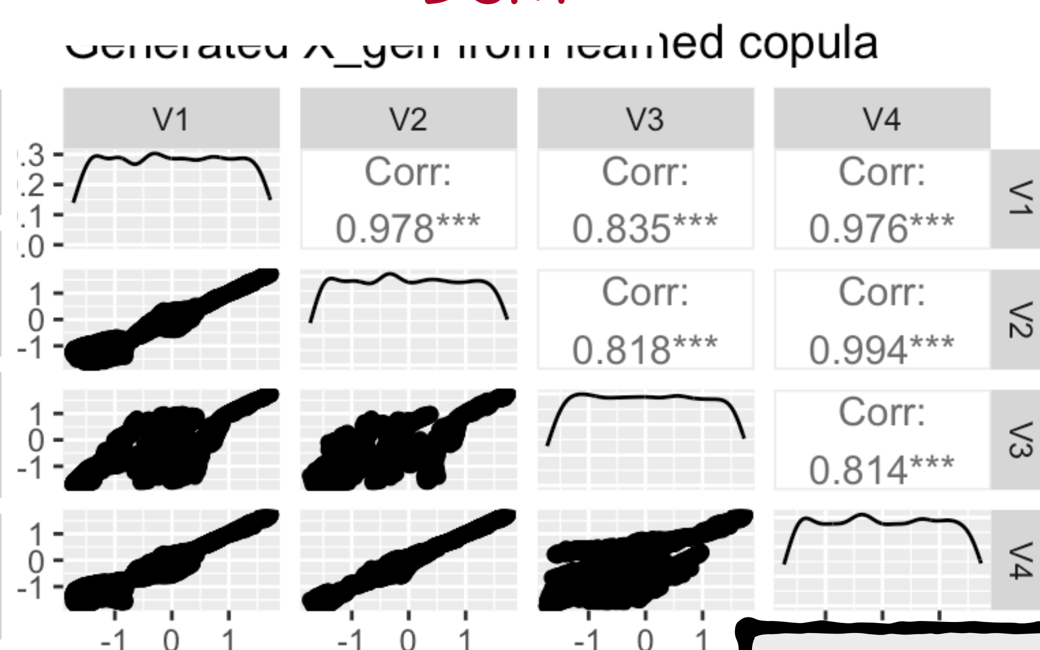
(Using DGMs to learn a non-parametric model)

$\{\xi_1, \ldots \xi_n\}$ → Construct Empirical Copula → $\{U_1, \ldots U_n\}$ → Learn Copula density using DGM → Output sample from DGM $(U_1^\star, \ldots, U_d^\star)$ → Output sample from target distribution $F_1^{-1}(U_1^\star), \ldots, F_d^{-1}(U_d^\star)$

Empirical Samples



DGM



The empirical copula: why this is a good approach?

(Learning joint distribution is hard!)

$\{\xi_1, \ldots \xi_n\}$ → Construct Empirical Copula → $\{U_1, \ldots U_n\}$ → Learn Copula density using DGM → Output sample from DGM $(U_1^\star, \ldots, U_d^\star)$ → Output sample from target distribution $F_1^{-1}(U_1^\star), \ldots, F_d^{-1}(U_d^\star)$

<u>Advantages of this construction</u>

- Only univariate marginals need to be modelled (parametric or empirical); no parametric joint model required.
- Flexible DGM copula model captures complex dependence reduces risk of misrepresenting dependence structure

Agenda for the course

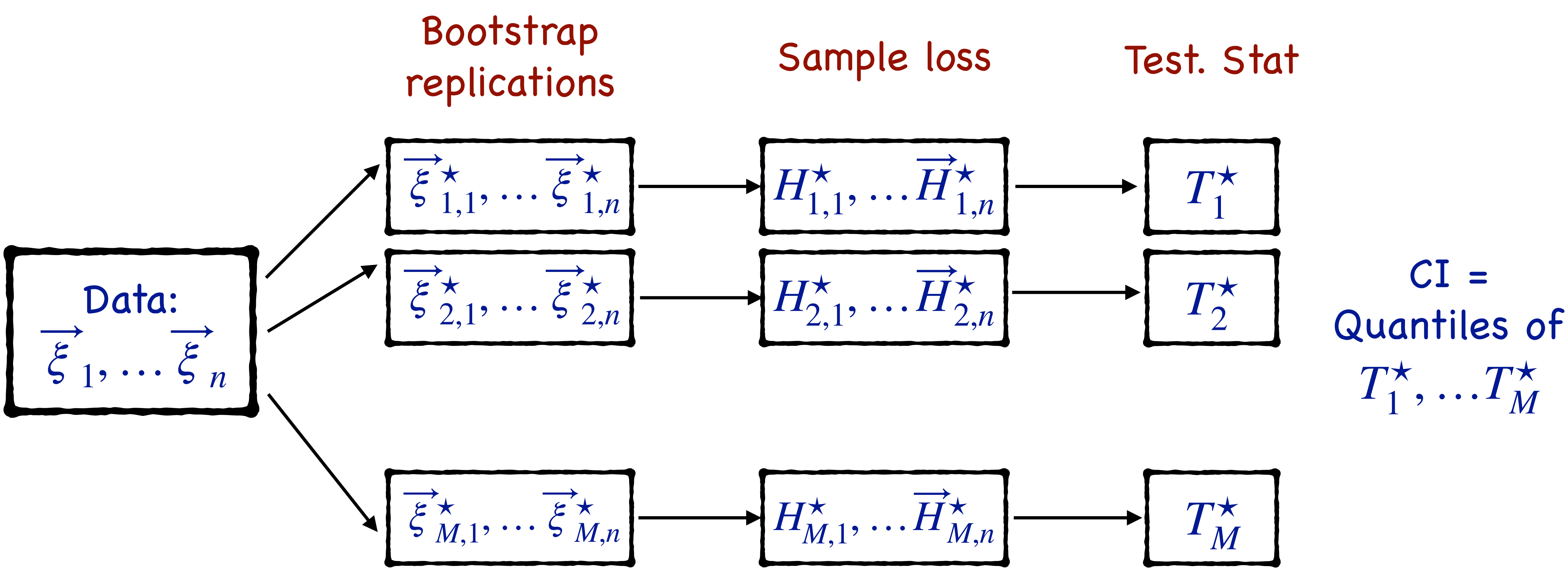
Session 1 Input Modelling 🚂	Session 2 Model Risk & Output Analysis 🚂
The bootstrap	The copula
Inverse Transforms	Model Risk
Sampling from multivariate distributions	

Model Risk: how errors in input modelling propagate

(Why good distributional assumption are essential)

Problem: The loss incurred upon investment in a complex financial instrument is given by $H = \varphi(x, \xi)$, where ξ are one day losses and x is an investment strategy. You are required to find the capital that needs to be set aside to that the probability of loss exceeding this amount is below a threshold. You are also required to find an investment strategy that minimises a measure $\rho(x, P)$ that is a function of the data distribution P and decision x

Approach 1: Use Bootstrap (Notation: $H = \varphi(\overrightarrow{\xi})$)

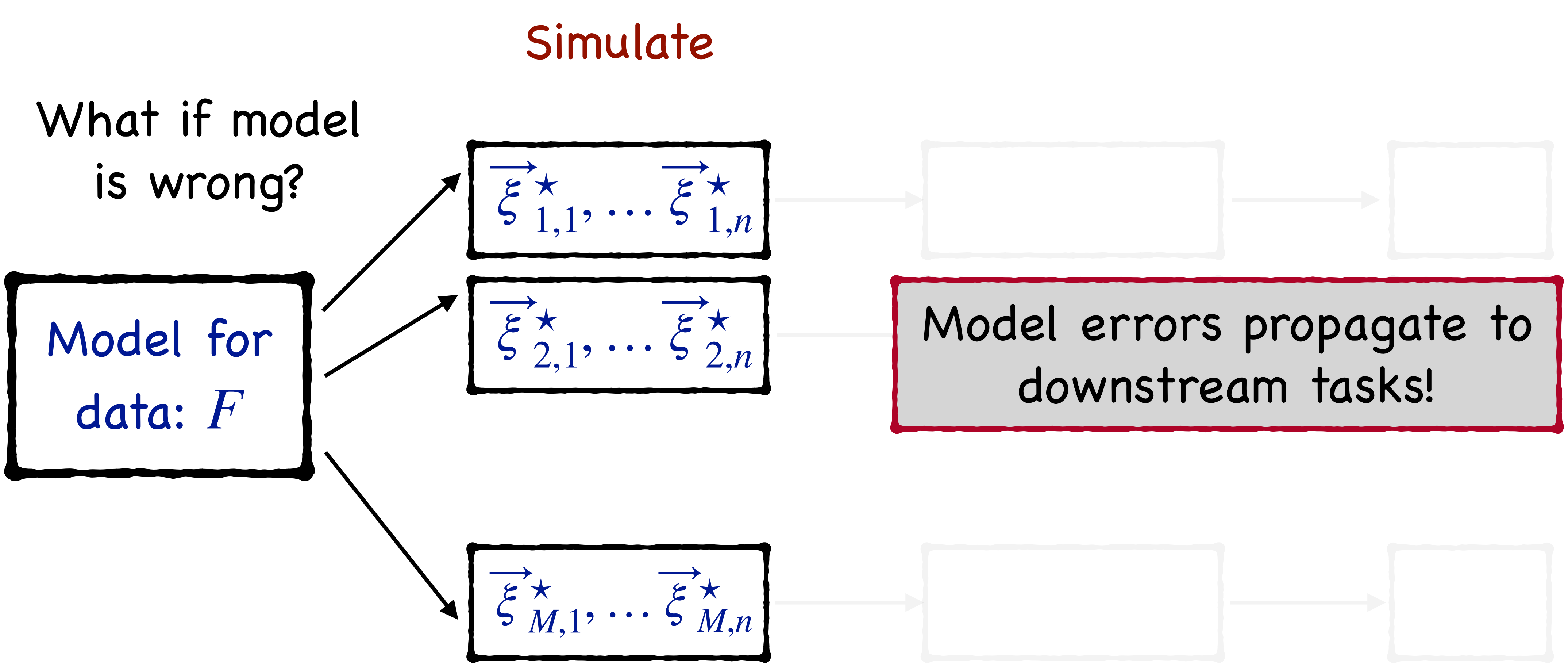


Model Risk: how errors in input modelling propagate

(Why good distributional assumption are essential)

<u>Problem:</u> The loss incurred upon investment in a complex financial instrument is given by $H = \varphi(x, \xi)$, where ξ are one day losses and x is an investment strategy. You are required to find the capital that needs to be set aside to that the probability of loss exceeding this amount is below a threshold. You are also required to find an investment strategy that minimises a measure $\rho(x, P)$ that is a function of the data distribution P and decision x

Approach 2: Assume a distribution for data (Notation: $H = \varphi(\vec{\xi})$)



Model Risk: how errors in input modelling propagate

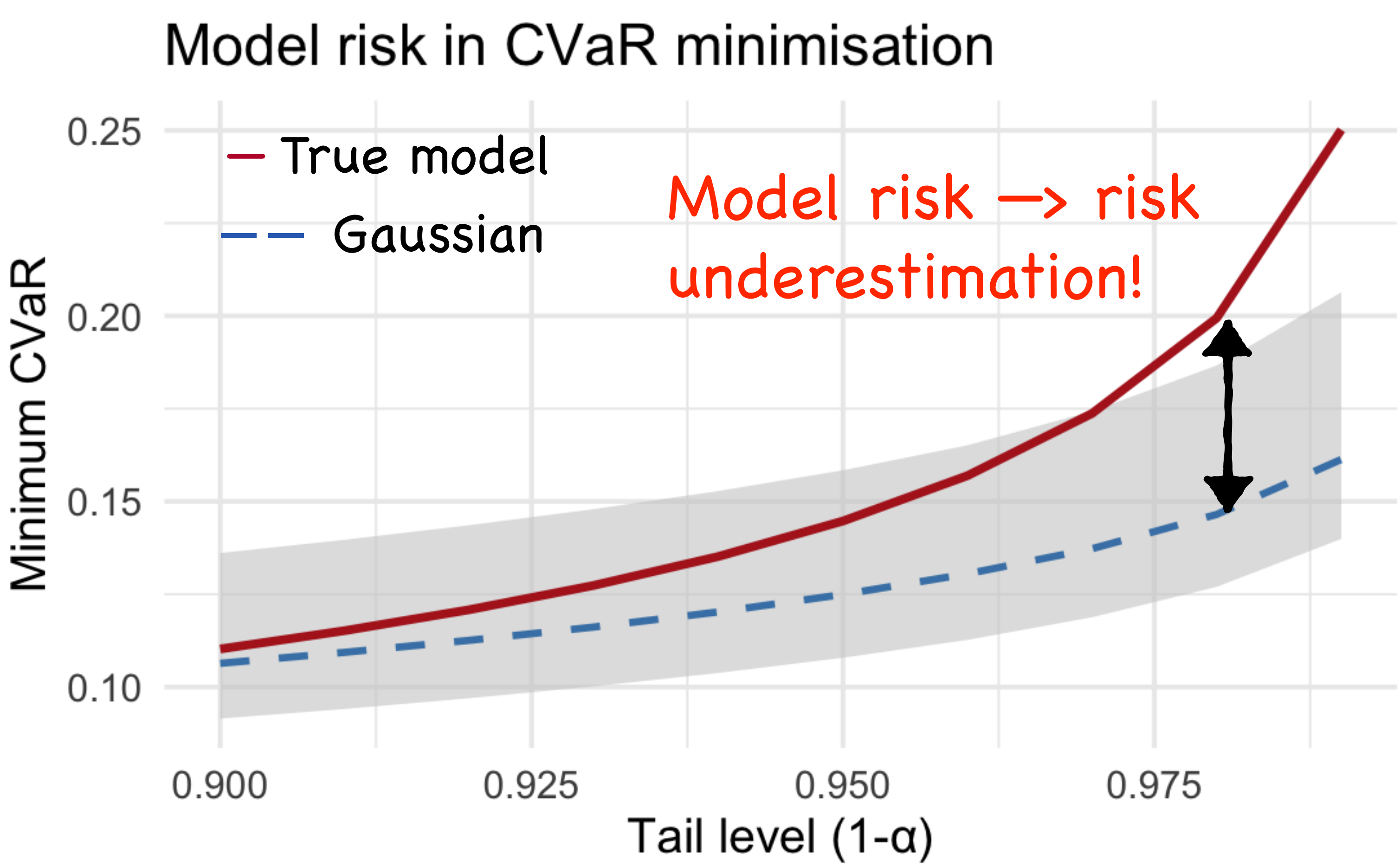
(Numerical Evidence)

Problem: Find the investment strategy that minimises

$$\rho(x, P) = \text{CVaR}_{1-\alpha}(x, P)$$

where $\text{CVaR}_{1-\alpha}(x, P) = E_P[\varphi(x, \xi) \mid \varphi(x, \xi) \geq \text{VaR}_{1-\alpha}(x)]$

Attempt: Given data samples $\{\xi_1, \ldots, \xi_n\}$ fit a Gaussian distribution to data. Minimise risk functional of interest assuming the model was Gaussian



Solve optimisation problem

$\min_x \rho(x, \hat{P})$ where $\hat{P} \rightarrow$ Gaussian

- $\hat{P} \rightarrow$ regulator proposed model
- Model risk —> risk due to wrong model choice.

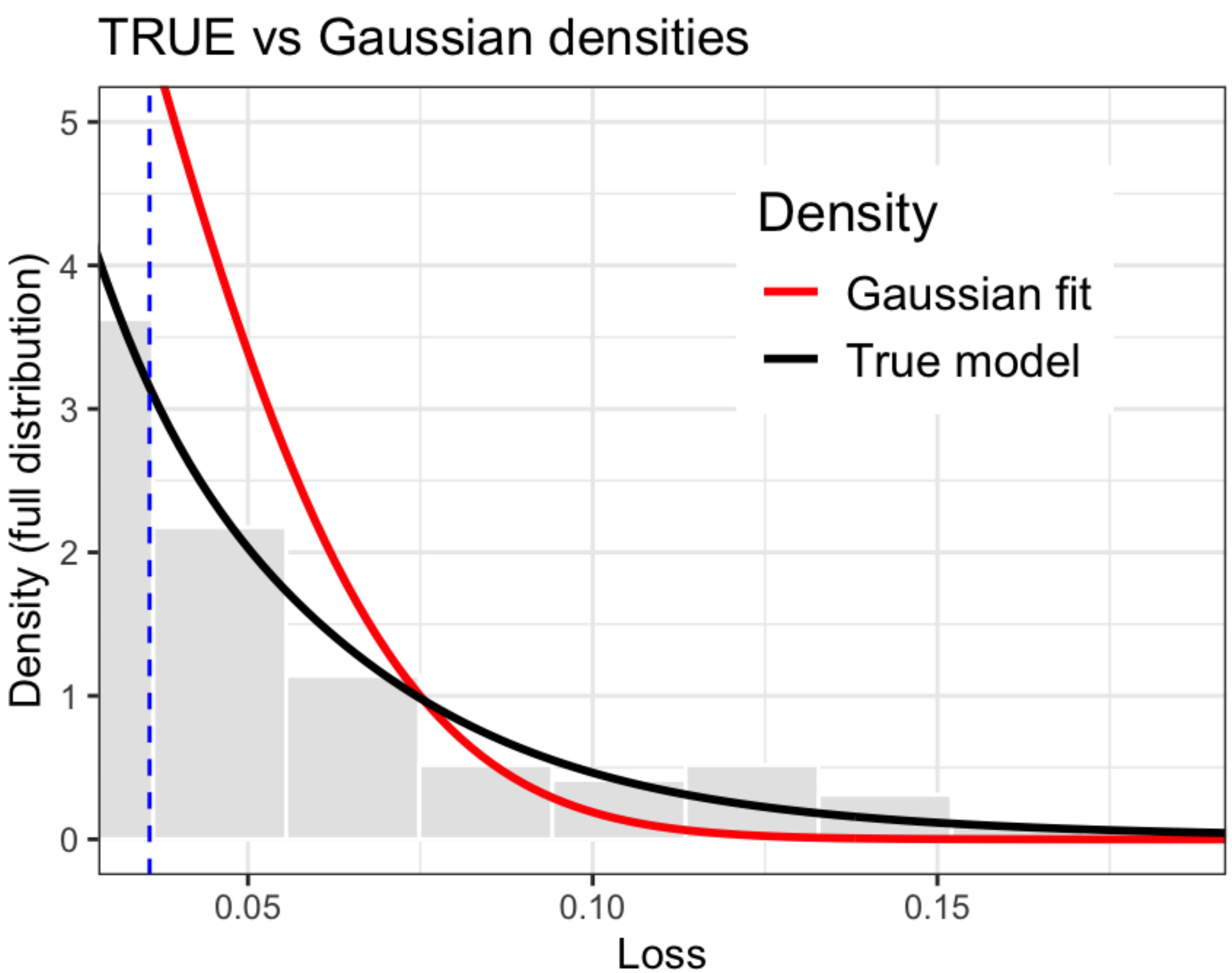
Model Risk: how errors in input modelling propagate

(Intuition)

Problem: Find the investment strategy that minimises

$$\rho(x, P) = \text{CVaR}_{1-\alpha}(x, P)$$

where $\text{CVaR}_{1-\alpha}(x, P) = E_P[\varphi(x, \xi) \mid \varphi(x, \xi) \geq \text{VaR}_{1-\alpha}(x)]$



- Gaussian distribution underestimates outliers in data (Model error)
- Leads to underestimation of risk (percolation of error)
- Alternative: sample distribution + bootstrap
- Catch: Paucity of tail samples $\Longrightarrow$ fragility to data realisation

Samples relevant to risk evaluation are mis-represented in model!

End of Lecture 2