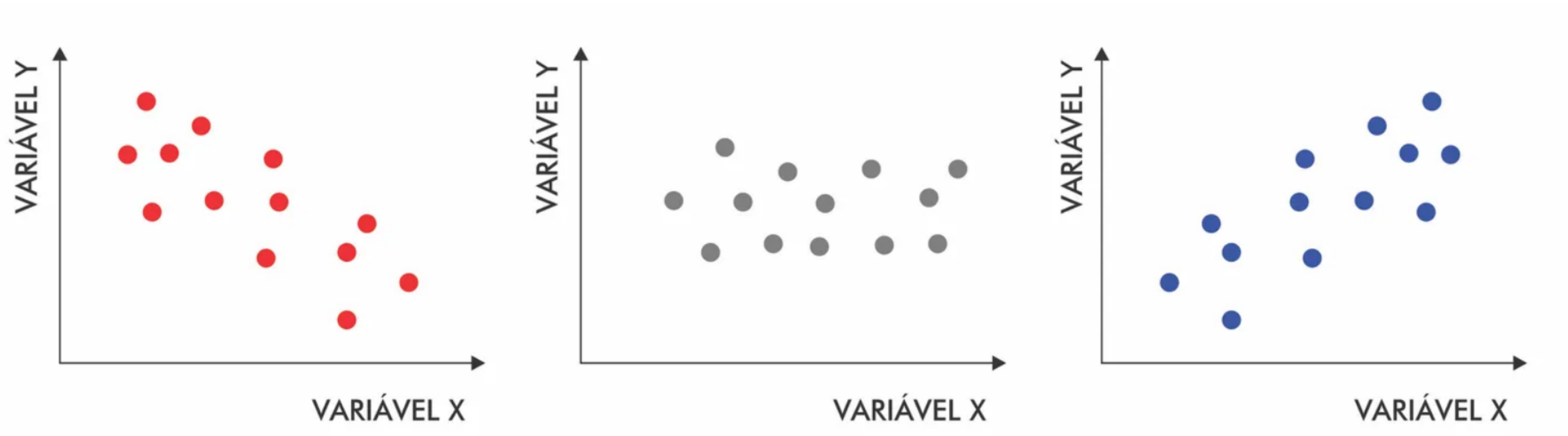
Variáveis correlacionadas

Correlação Linear

- Correlação $\Rightarrow$ dependência, mas dependência $\nRightarrow$ correlação

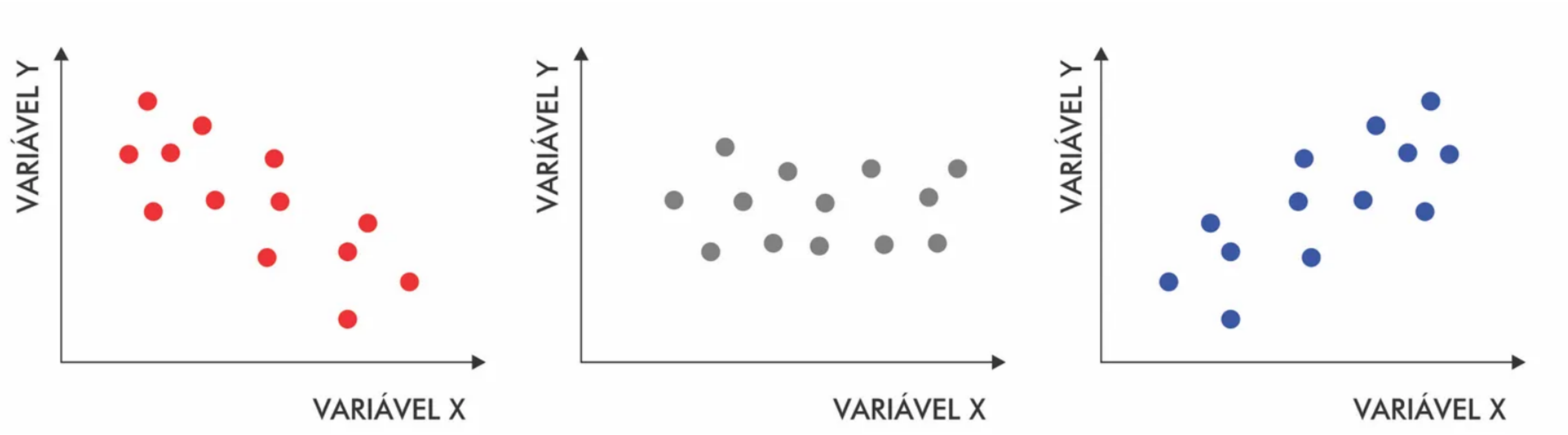
Correlação Linear

- Correlação ⇒ dependência, mas dependência ⇏ correlação



Correlação Linear

- Correlação ⇒ dependência, mas dependência ⇏ correlação



$$\rho_{XY} = \frac{Cov[X,Y]}{\sqrt{Var[X]}\sqrt{Var[Y]}} = \frac{X \cdot Y}{\|X\| * \|Y\|} = \cos \alpha_{XY}$$

Teoria de Cópulas

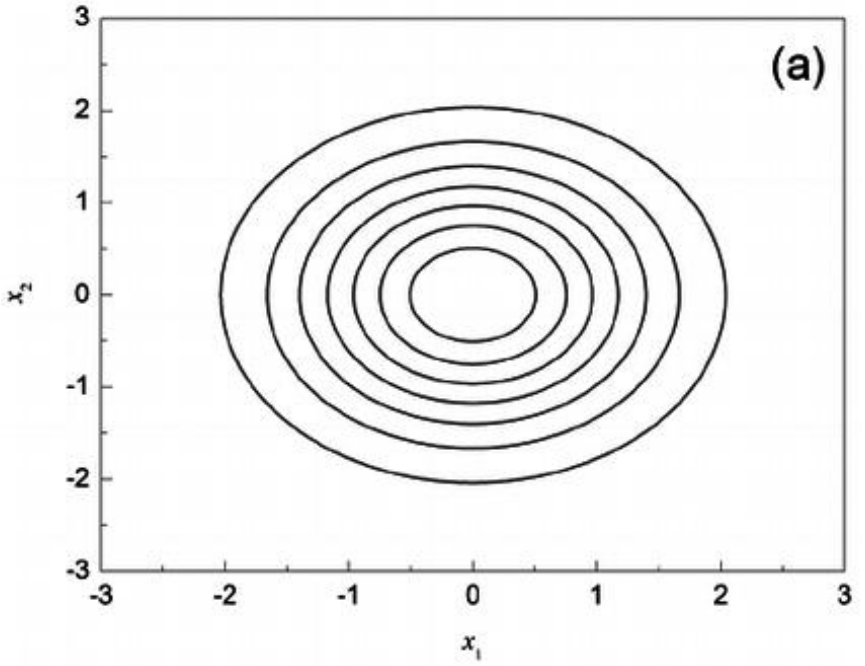
- Estrutura de dependência;

- Mais genérica do que a correlação (que se resume a um único tipo de cópula);

- Bem como consiste em um demétodo que cria estrutura de pêndencias entre distribuições marginais, a fim de criar uma distribuição conjunta.

Tipo de Cópulas

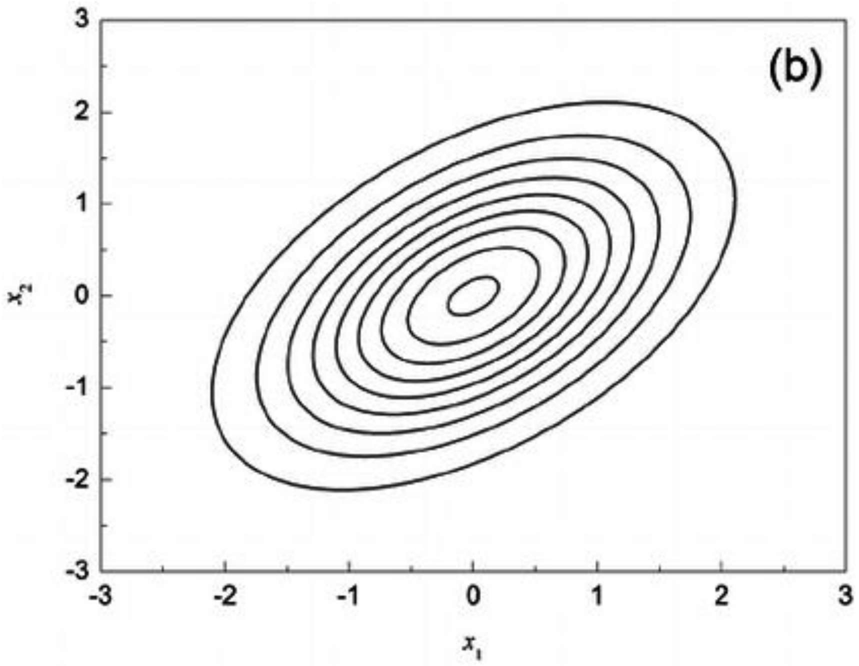
- **Cópula Independente;**



Tipo de Cópulas

- Cópula Independente;

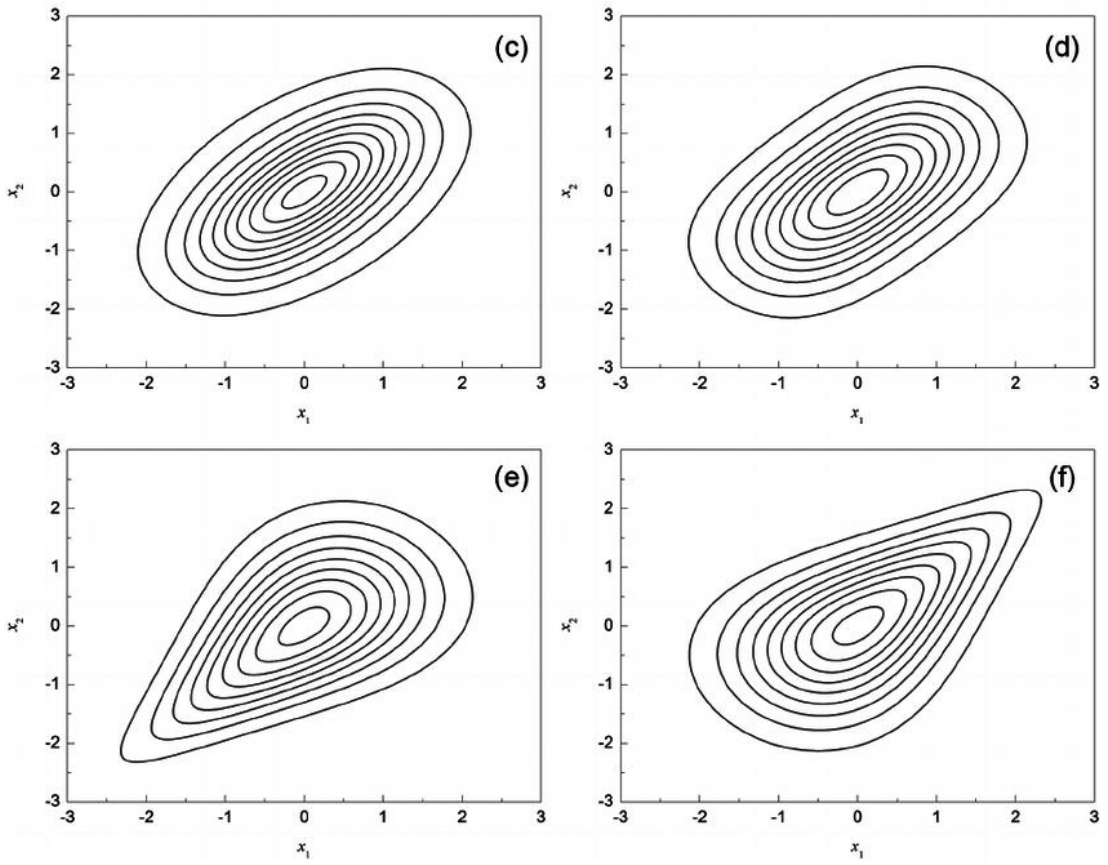
- **Cópula gaussiana;**



EXEMPLO - MATLAB

Tipo de Cópulas

- Cópula Independente;
- Cópula gaussiana;
- Cópula de Plackett (c);
- Cópula de Frank (d);
- Cópula de Clayton (e – f);



PCE – Para variáveis não independentes

- A teoria do PCE indica a necessidade de variáveis independentes. Então, como lidar quando as variáveis não são independentes?

(i) Ignorar a correlação das variáveis;

(ii) Através de transformações isoprobabilísticas, transformar duas variáveis dependentes em duas variáveis independentes;

(iii) Construir um PCE com polinômios arbitrários, que sejam ortogonais às distribuições marginais do problema.

PCE – Para variáveis não independentes

- A teoria do PCE indica a necessidade de variáveis independentes. Então, como lidar quando as variáveis não são independentes?

(i) Ignorar a correlação das variáveis;

(ii) Através de transformações isoprobabilísticas, transformar duas variáveis dependentes em duas variáveis independentes;

(iii) Construir um PCE com polinômios arbitrários, que sejam ortogonais às distribuições marginais do problema.

PCE – Para variáveis não independentes

- A teoria do PCE indica a necessidade de variáveis independentes. Então, como lidar quando as variáveis não são independentes?

(i) **Ignorar a correlação das variáveis;**

(ii) **Através de transformações isoprobabilísticas, transformar duas variáveis dependentes em duas variáveis independentes;**

(iii) Construir um PCE com polinômios arbitrários, que sejam ortogonais às distribuições marginais do problema.

Métodos Globais

- Método de Morris
- Índices de Borgonovo
- Índices Sobol
- Índices ANCOVA
- Índices Kucherenko
- Método de Cotter*

Índices Sobol

$$\boldsymbol{Y} = \mathcal{M}(\boldsymbol{X})$$

Supõe que o modelo pode ser decomposto em HDMR

$$\boldsymbol{Y} = \mathcal{M}_0 + \sum_{i=1}^{M} \mathcal{M}_i(x_i) + \sum_{1 \leq i \leq j \leq M}^{M} \mathcal{M}_{ij}(x_i, x_j) + \cdots + \mathcal{M}_{1,2,\ldots,M}(x_1, x_2, \ldots, x_M)$$

ANCOVA – Análise de Covariância

- Li e Rabitz (2010) indicam que é possível a decomposição de variâncias para variáveis correlacionadas. No entanto, a variância total deixa de ser escrita em parcelas de variâncias, e passa a ser descrita por relações de covariância.

ANCOVA – Análise de Covariância

- Li e Rabitz (2010) indicam que é possível a decomposição de variâncias para variáveis correlacionadas. No entanto, a variância total deixa de ser escrita em parcelas de variâncias, e passa a ser descrita por relações de covariância.
- A partir desse HDMR, os índices ANCOVA são definidos da seguinte maneira:

$$S_i = \frac{Cov[\mathcal{M}_i(x_i), Y]}{Var[Y]} \Rightarrow S_i = \frac{Var[\mathcal{M}_i(x_i)]}{Var[Y]} + \frac{Cov[\mathcal{M}_i(x_i), (Y - \mathcal{M}_i(x_i))]}{Var[Y]}$$

ANCOVA – Análise de Covariância

- Li e Rabitz (2010) indicam que é possível a decomposição de variâncias para variáveis correlacionadas. No entanto, a variância total deixa de ser escrita em parcelas de variâncias, e passa a ser descrita por relações de covariância.

- A partir desse HDMR, os índices ANCOVA são definidos da seguinte maneira:

$$S_i = \frac{Cov[\mathcal{M}_i(x_i), Y]}{Var[Y]} \Rightarrow S_i = \underbrace{\boxed{\frac{Var[\mathcal{M}_i(x_i)]}{Var[Y]}}}_{S_i^U} + \underbrace{\boxed{\frac{Cov[\mathcal{M}_i(x_i), (Y - \mathcal{M}_i(x_i))]}{Var[Y]}}}_{S_i^C}$$

ANCOVA – Análise de Covariância

- Li e Rabitz (2010) indicam que é possível a decomposição de variâncias para variáveis correlacionadas. No entanto, a variância total deixa de ser escrita em parcelas de variâncias, e passa a ser descrita por relações de covariância.

- A partir desse HDMR, os índices ANCOVA são definidos da seguinte maneira:

$$S_i = \frac{Cov[\mathcal{M}_i(x_i), Y]}{Var[Y]} \Rightarrow S_i = \underbrace{\boxed{\frac{Var[\mathcal{M}_i(x_i)]}{Var[Y]}}}_{S_i^U} + \underbrace{\boxed{\frac{Cov[\mathcal{M}_i(x_i), (Y - \mathcal{M}_i(x_i))]}{Var[Y]}}}_{S_i^C}$$

- Caniou (2012) sugere o uso de PCEs para a construção do HDMR no cálculo de índices ANCOVA.

ANCOVA – Análise de Covariância

- Os índices totais ficariam:

$$S_i^{U,T} = \frac{Var[\mathcal{M}_{i\in u}(x_i)]}{Var[Y]}$$

$$S_i^{C,T} = \frac{Cov[\mathcal{M}_{i\in u}(x), \mathcal{M}_{i\notin v}(x)]}{Var[Y]}$$

ANCOVA – Análise de Covariância

- Os índices totais ficariam:

$$S_i^{U,T} = \frac{Var[\mathcal{M}_{\text{[illegible]}}(x_i)]}{\text{[illegible]}]}$$

$$S_i^{C,T} = \frac{Co\text{[illegible]}\in u\text{[illegible]}\mathcal{M}_{i\notin v}(x)]}{Var[Y]}$$

ANCOVA – Análise de Covariância

- Os índices totais ficariam:

$$\cancel{S_i^{U,T} = \frac{Var[\mathcal{M}_{i\in u}(x_i)]}{Var[Y]}}$$

$$\cancel{S_i^{C,T} = \frac{Cov[\mathcal{M}_{i\in u}(x), \mathcal{M}_{i\notin v}(x)]}{Var[Y]}}$$

$$S_i = S_i^U + S_i^I + S_i^C = \frac{Var[\mathcal{M}_i(x_i)]}{Var[Y]} + \frac{Cov[\mathcal{M}_i(x_i), \mathcal{M}_{i\in u}(x)]}{Var[Y]} + \frac{Cov[\mathcal{M}_i(x_i), \mathcal{M}_{i\notin v}(x)]}{Var[Y]}$$

ANCOVA – Análise de Covariância

- Os índices totais ficariam:

$$\cancel{S_i^{U,T} = \frac{Var[\mathcal{M}_{[illegible]}(x_i)]}{[illegible]]}}$$

$$\cancel{S_i^{C,T} = \frac{Co[illegible]\in u[illegible]\mathcal{M}_{i\notin v}(x)]}{Var[Y]}}$$

$$S_i = S_i^U + S_i^I + S_i^C = \frac{Var[\mathcal{M}_i(x_i)]}{Var[Y]} + \frac{Cov[\mathcal{M}_i(x_i), \mathcal{M}_{i\in u}(x)]}{Var[Y]} + \frac{Cov[\mathcal{M}_i(x_i), \mathcal{M}_{i\notin v}(x)]}{Var[Y]}$$

Métodos Globais

- Método de Morris
- Índices de Borgonovo
- Índices Sobol
- Índices ANCOVA
- Índices Kucherenko
- Método de Cotter*

Método de Kucherenko

- Utiliza a lei da variância total para definir uma nova metodologia de partição da variância total. Isso permite escrever a variância total como:

$$Var[Y] = Var_i\left[E_{\sim i}[Y|X_i]\right] + E_i[Var_{\sim i}[Y|X_i]]$$

$$1 = \underbrace{\boxed{\frac{Var_i\left[E_{\sim i}[Y|X_i]\right]}{Var[Y]}}}_{S_i} + \underbrace{\boxed{\frac{E_i[Var_{\sim i}[Y|X_i]]}{Var[Y]}}}_{S_{\sim i}^T}$$

Método de Kucherenko

$$S_i^T = \frac{E_{\sim i}[Var_i[Y|X_i]]}{Var[Y]}$$

$$S_i = \frac{Var_i\left[E_{\sim i}[Y|X_i]\right]}{Var[Y]}$$

Estes valores podem ser estimados através de um loop duplo de Monte Carlo!