Copula-based robust optimal block designs

Werner G. Müller, joint work with Dave Woods

based on joint work with E. Perrone and A. Rappold at

IFAS - Institut für Angewandte Statistik, JKU Linz, Austria

Marseille, May 2^{nd} 2018

JOHANNES KEPLER UNIVERSITY LINZ

What is a copula function?

and what is it good for?

Definition (Copula)

Let $\mathbb{I} = [0,1]$. A *two-dimensional copula* (or *2-copula*) is a bivariate function $\mathbf{C} : \mathbb{I} \times \mathbb{I} \longrightarrow \mathbb{I}$ with the following properties:

1. for every u, $v \in \mathbb{I}$

$$\mathbf{C}(u,0) = 0,\ \mathbf{C}(u,1) = u,\ \mathbf{C}(0,v) = 0,\ \mathbf{C}(1,v) = v; \tag{1}$$

2. for every u_1, v_1, u_2, $v_2 \in \mathbb{I}$ such that $u_1 \leq u_2$ and $v_1 \leq v_2$,

$$\mathbf{C}(u_2,v_2) - \mathbf{C}(u_2,v_1) - \mathbf{C}(u_1,v_2) + \mathbf{C}(u_1,v_1) \geq 0. \tag{2}$$

JOHANNES KEPLER UNIVERSITY LINZ

Sklar's Theorem

bivariate case

Theorem (Sklar's Theorem)

Let $\mathbf{F}_{Y_1 Y_2}$ be a joint distribution function with marginals F_{Y_1} and F_{Y_2}. Then there exists a 2-copula $\mathbf{C}$ such that

$$\mathbf{F}_{Y_1 Y_2}(y_1, y_2) = \mathbf{C}(F_{Y_1}(y_1), F_{Y_2}(y_2)) \tag{3}$$

for all reals y_1, y_2.
If F_{Y_1} and F_{Y_2} are continuous, then $\mathbf{C}$ is unique.

Conversely, if $\mathbf{C}$ is a 2-copula and F_{Y_1} and F_{Y_2} are distribution functions, then the function $F_{Y_1 Y_2}$ given by (3) is a joint distribution with marginals F_{Y_1} and F_{Y_2}.

Sklar's Theorem

bivariate case

Theorem (Sklar's Theorem)

Let $\mathbf{F}_{Y_1 Y_2}$ be a joint distribution function with marginals F_{Y_1} and F_{Y_2}. Then there exists a 2-copula $\mathbf{C}$ such that

$$\mathbf{F}_{Y_1 Y_2}(y_1, y_2) = \mathbf{C}(F_{Y_1}(y_1), F_{Y_2}(y_2)) \tag{3}$$

for all reals y_1, y_2.
If F_{Y_1} and F_{Y_2} are continuous, then $\mathbf{C}$ is unique.

Conversely, if $\mathbf{C}$ is a 2-copula and F_{Y_1} and F_{Y_2} are distribution functions, then the function $F_{Y_1 Y_2}$ given by (3) is a joint distribution with marginals F_{Y_1} and F_{Y_2}.

JOHANNES KEPLER UNIVERSITY LINZ

A regression model

based on copulas

Let us consider a vector $\mathbf{x}^T = (x_1, \ldots, x_r) \in \mathcal{X}$ of control variables, where $\mathcal{X} \subset \mathbb{R}^r$ is a compact set.
The result of the observations is the vector:

$$\mathbf{y}(\mathbf{x}) = (y_1(\mathbf{x}), , \ldots, y_m(\mathbf{x})),$$

with

$$\mathbf{E}[\mathbf{Y}(x)] = \eta(\mathbf{x}, \beta) = (\eta_1(\mathbf{x}, \beta), \ldots, \eta_m(\mathbf{x}, \beta)),$$

where $\beta = (\beta_1, \ldots, \beta_k)$ is a certain unknown parameter vector to be estimated and η_i are known functions.

In our examples we focus on the case $m = 2$.

Define $c_{\mathbf{Y}}(\mathbf{y}(\mathbf{x}, \beta), \alpha)$ the joint probability density function of the random vector $\mathbf{Y}$, where $\alpha = (\alpha_1, \ldots, \alpha_l)$ are unknown parameters.

JOHANNES KEPLER UNIVERSITY LINZ

A regression model

based on copulas

Let us consider a vector $\mathbf{x}^T = (x_1, \ldots, x_r) \in \mathcal{X}$ of control variables, where $\mathcal{X} \subset \mathbb{R}^r$ is a compact set.
The result of the observations is the vector:

$$\mathbf{y}(\mathbf{x}) = (y_1(\mathbf{x}), , \ldots, y_m(\mathbf{x})),$$

with

$$\mathbf{E}[\mathbf{Y}(x)] = \eta(\mathbf{x}, \beta) = (\eta_1(\mathbf{x}, \beta), \ldots, \eta_m(\mathbf{x}, \beta)),$$

where $\beta = (\beta_1, \ldots, \beta_k)$ is a certain unknown parameter vector to be estimated and η_i are known functions.

In our examples we focus on the case $m = 2$.

Define $c_{\mathbf{Y}}(\mathbf{y}(\mathbf{x}, \beta), \alpha)$ the joint probability density function of the random vector $\mathbf{Y}$, where $\alpha = (\alpha_1, \ldots, \alpha_l)$ are unknown parameters.

JOHANNES KEPLER UNIVERSITY LINZ

A regression model

based on copulas

Let us consider a vector $\mathbf{x}^T = (x_1, \dots, x_r) \in \mathcal{X}$ of control variables, where $\mathcal{X} \subset \mathbb{R}^r$ is a compact set.
The result of the observations is the vector:

$$\mathbf{y}(\mathbf{x}) = (y_1(\mathbf{x}), , \dots, y_m(\mathbf{x})),$$

with

$$\mathbf{E}[\mathbf{Y}(x)] = \eta(\mathbf{x}, \beta) = (\eta_1(\mathbf{x}, \beta), \dots, \eta_m(\mathbf{x}, \beta)),$$

where $\beta = (\beta_1, \dots, \beta_k)$ is a certain unknown parameter vector to be estimated and η_i are known functions.

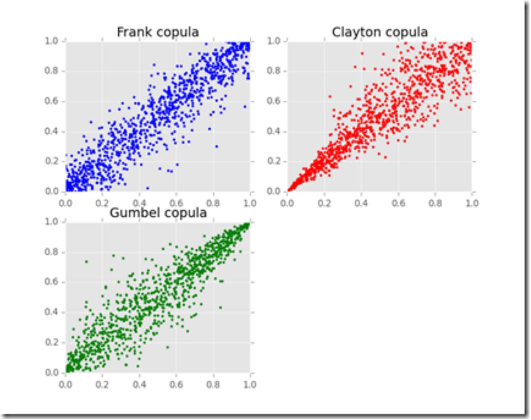
In our examples we focus on the case $m = 2$.

Define $c_{\mathbf{Y}}(\mathbf{y}(\mathbf{x}, \beta), \alpha)$ the joint probability density function of the random vector $\mathbf{Y}$, where $\alpha = (\alpha_1, \dots, \alpha_l)$ are unknown parameters.

JKU
JOHANNES KEPLER
UNIVERSITY LINZ

Scatterplots of typical copulas

from the internet



Used copula functions

connect through Kendall's τ_2

1. *Product Copula,* which represents the independence case:

$$C(u_1, u_2) = u_1 u_2,$$

   with $\tau_2 = 0$.

2. *Clayton Copula:*

$$C_\alpha(u_1, u_2) = \left[\max\left(u_1^{-\alpha} + u_2^{-\alpha} - 1, 0\right)\right]^{-\frac{1}{\alpha}},$$

   with $\alpha \in (0, +\infty)$ and $\tau_2 = \frac{\alpha}{\alpha+2}$.

3. *Gumbel Copula:*

$$C_\alpha(u_1, u_2) = \exp\left(-\left[(-\ln u_1)^\alpha + (-\ln u_2)^\alpha\right]^{\frac{1}{\alpha}}\right),$$

   with $\alpha \in [1, +\infty)$ and $\tau_2 = \frac{\alpha-1}{\alpha}$.

JOHANNES KEPLER UNIVERSITY LINZ

The approach

completing the model

Remember

according to Sklar's theorem the joint probability density function is the density of the copula function such that

$$F_{Y_1,Y_2}(y_1,y_2;\alpha) = \int c_{\mathbf{Y}}(\mathbf{y}(\mathbf{x},\beta),\alpha)d\mathbf{y} =$$

$$= C(F_{Y_1}(y_1),F_{Y_2}(y_2);\alpha)$$

For r independent observations at $x_1,\ldots,x_r$, the corresponding Information matrix is

$$\mathbf{M}(\xi,\beta,\alpha) = \sum_{i=1}^{r} w_i J(x_i,\beta,\alpha)$$

JOHANNES KEPLER UNIVERSITY LINZ

The approach

completing the model

Remember

according to Sklar's theorem the joint probability density function is the density of the copula function such that

$$F_{Y_1,Y_2}(y_1,y_2;\alpha) = \int c_{\mathbf{Y}}(\mathbf{y}(\mathbf{x},\beta),\alpha)d\mathbf{y} =$$

$$= C(F_{Y_1}(y_1),F_{Y_2}(y_2);\alpha)$$

For r independent observations at $x_1,\ldots,x_r$, the corresponding Information matrix is

$$\mathbf{M}(\xi,\beta,\alpha) = \sum_{i=1}^{r} w_i J(x_i,\beta,\alpha)$$

JOHANNES KEPLER UNIVERSITY LINZ

The Fisher information

for all parameters

Definition

For a single observation the matrix $J(\mathbf{x}, \beta, \alpha)$, a $(k+l) \times (k+l)$ matrix defined as follows

$$J(\mathbf{x}, \beta, \alpha) = \begin{pmatrix} J_{\beta\beta}(\mathbf{x}) & J_{\beta\alpha}(\mathbf{x}) \\ J_{\beta\alpha}^T(\mathbf{x}) & J_{\alpha\alpha}(\mathbf{x}) \end{pmatrix} \quad (4)$$

where the matrix $J_{\beta\beta}(\mathbf{x})$ is the $(k \times k)$ matrix with the (i,j)th element defined as

$$\mathbf{E}\left(-\frac{\partial^2}{\partial\beta_i \partial\beta_j} \log c_{\mathbf{Y}}(\mathbf{y}(\mathbf{x}, \beta), \alpha)\right) = \quad (5)$$

$$= \mathbf{E}\left(\left(\frac{\partial}{\partial\beta} \log c_{\mathbf{Y}}(\mathbf{y}(\mathbf{x}, \beta), \alpha)\right)\left(\frac{\partial}{\partial\beta} \log c_{\mathbf{Y}}(\mathbf{y}(\mathbf{x}, \beta), \alpha)\right)^T\right)$$

and so are also the matrices $J_{\beta\alpha}(\mathbf{x})$ and $J_{\alpha\alpha}(\mathbf{x})$, respectively.

Optimal design

basics

Design
$$\xi = \left\{ \begin{array}{cccc} \mathbf{x_1} & \mathbf{x_2} & \dots & \mathbf{x_n} \\ w_1 & w_2 & \dots & w_n \end{array} \right\},$$

where $\sum_{i=1}^{r} w_i = 1$.

Aim

We are concerned with finding $\xi^*(\beta, \alpha)$ such that maximizes some scalar function $\phi(M(\xi, \beta, \alpha))$.

D-optimality

For now we consider the function $\phi(M) = \log\det M$, i.e. *D-optimality*.

JOHANNES KEPLER UNIVERSITY LINZ

Optimal design

basics

Design
$$\xi = \left\{ \begin{array}{cccc} \mathbf{x_1} & \mathbf{x_2} & \dots & \mathbf{x_n} \\ w_1 & w_2 & \dots & w_n \end{array} \right\},$$

where $\sum_{i=1}^{r} w_i = 1$.

Aim

We are concerned with finding $\xi^*(\beta,\alpha)$ such that maximizes some scalar function $\phi(M(\xi,\beta,\alpha))$.

D-optimality

For now we consider the function $\phi(M) = \log\det M$, i.e. *D-optimality*.

Optimal design

basics

Design
$$\xi = \left\{ \begin{array}{cccc} \mathbf{x_1} & \mathbf{x_2} & \dots & \mathbf{x_n} \\ w_1 & w_2 & \dots & w_n \end{array} \right\},$$

where $\sum_{i=1}^{r} w_i = 1$.

Aim

We are concerned with finding $\xi^*(\beta,\alpha)$ such that maximizes some scalar function $\phi(M(\xi,\beta,\alpha))$.

D-optimality

For now we consider the function $\phi(M) = \log\det M$, i.e. *D-optimality*.

JOHANNES KEPLER UNIVERSITY LINZ

Design theory

Kiefer and Wolfowitz type, cf. Heise and Myers, 1996

Theorem (Equivalence Theorem: D-criterion)

For a localized parameter vector $(\bar{\beta}, \bar{\alpha})$, the following properties are equivalent:

1. ξ^* *is D-optimal;*
2. $\text{tr}\,[M(\xi^*, \bar{\beta}, \bar{\alpha})^{-1} J(x, \bar{\beta}, \bar{\alpha})] \leq (k+l)$, $\forall x \in \mathcal{X}$;
3. ξ^* *minimize* $\max_{x \in \mathcal{X}} \text{tr}\,[M(\xi^*, \bar{\beta}, \bar{\alpha})^{-1} J(x, \bar{\beta}, \bar{\alpha})]$, *over all* $\xi \in \Xi$.

The function $d(x, \xi^*) = \text{tr}\,[M(\xi^*, \bar{\beta}, \bar{\alpha})^{-1} J(x, \bar{\beta}, \bar{\alpha})]$ is the *sensitivity function* for the D-criterion.

Two designs ξ, ξ^* can be compared by the following ratio (*D-Efficiency*):

$$\left(\frac{|M(\xi, \beta, \alpha)|}{|M(\xi^*, \beta, \alpha)|} \right)^{1/(k+l)}.$$

Design theory

Kiefer and Wolfowitz type, cf. Heise and Myers, 1996

Theorem (Equivalence Theorem: D-criterion)

For a localized parameter vector $(\bar{\beta}, \bar{\alpha})$, the following properties are equivalent:

1. ξ^* *is D-optimal;*
2. $\text{tr}\,[M(\xi^*, \bar{\beta}, \bar{\alpha})^{-1} J(x, \bar{\beta}, \bar{\alpha})] \leq (k+l)$, $\forall x \in \mathcal{X}$;
3. ξ^* *minimize* $\max_{x \in \mathcal{X}} \text{tr}\,[M(\xi^*, \bar{\beta}, \bar{\alpha})^{-1} J(x, \bar{\beta}, \bar{\alpha})]$, *over all* $\xi \in \Xi$.

The function $d(x, \xi^*) = \text{tr}\,[M(\xi^*, \bar{\beta}, \bar{\alpha})^{-1} J(x, \bar{\beta}, \bar{\alpha})]$ is the *sensitivity function* for the D-criterion.

Two designs ξ, ξ^* can be compared by the following ratio (*D-Efficiency*):

$$\boxed{\left(\frac{|M(\xi, \beta, \alpha)|}{|M(\xi^*, \beta, \alpha)|} \right)^{1/(k+l)}.}$$

Design theory

Kiefer and Wolfowitz type, cf. Heise and Myers, 1996

Theorem (Equivalence Theorem: D-criterion)

For a localized parameter vector $(\bar{\beta}, \bar{\alpha})$, the following properties are equivalent:

1. ξ^* *is D-optimal;*
2. $\operatorname{tr}\,[M(\xi^*, \bar{\beta}, \bar{\alpha})^{-1} J(x, \bar{\beta}, \bar{\alpha})] \leq (k+l),\ \forall x \in \mathcal{X};$
3. ξ^* *minimize* $\max_{x \in \mathcal{X}} \operatorname{tr}\,[M(\xi^*, \bar{\beta}, \bar{\alpha})^{-1} J(x, \bar{\beta}, \bar{\alpha})]$, *over all* $\xi \in \Xi$.

The function $d(x, \xi^*) = \operatorname{tr}\,[M(\xi^*, \bar{\beta}, \bar{\alpha})^{-1} J(x, \bar{\beta}, \bar{\alpha})]$ is the *sensitivity function* for the D-criterion.
Two designs ξ, ξ^* can be compared by the following ratio (*D-Efficiency*):

$$\boxed{\left(\frac{|M(\xi, \beta, \alpha)|}{|M(\xi^*, \beta, \alpha)|} \right)^{1/(k+l)}.}$$

JOHANNES KEPLER UNIVERSITY LINZ

Another criterion

for parameter subsets

D_s-optimality

Another well known criterion is the D_s-criterion, i.e., the criterion $\phi_s(M) = \log\det(M_{11} - M_{12}M_{22}^{-1}M_{12}^T)$, if M is nonsingular and where

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{12}^T & M_{22} \end{pmatrix},$$

with M_{11} is the $(s \times s)$ minor related to the estimated parameters.

Definition

For the comparison of designs we define D_s-*Efficiency* of the design ξ with respect to the design ξ^* as the ratio

$$\boxed{\left(\frac{|M_{11}(\xi,\bar{\gamma}) - M_{12}(\xi,\bar{\gamma})M_{22}^{-1}(\xi,\bar{\gamma})M_{12}^T(\xi,\bar{\gamma})|}{|M_{11}(\xi^*,\bar{\gamma}) - M_{12}(\xi^*,\bar{\gamma})M_{22}^{-1}(\xi^*,\bar{\gamma})M_{12}^T(\xi^*,\bar{\gamma})|}\right)^{1/s}.}$$

Another criterion

for parameter subsets

D_s-optimality

Another well known criterion is the D_s-criterion, i.e., the criterion $\phi_s(M) = \log\det(M_{11} - M_{12}M_{22}^{-1}M_{12}^T)$, if M is nonsingular and where

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{12}^T & M_{22} \end{pmatrix},$$

with M_{11} is the $(s \times s)$ minor related to the estimated parameters.

Definition

For the comparison of designs we define D_s*-Efficiency* of the design ξ with respect to the design ξ^* as the ratio

$$\boxed{\left(\frac{|M_{11}(\xi,\bar{\gamma}) - M_{12}(\xi,\bar{\gamma})M_{22}^{-1}(\xi,\bar{\gamma})M_{12}^T(\xi,\bar{\gamma})|}{|M_{11}(\xi^*,\bar{\gamma}) - M_{12}(\xi^*,\bar{\gamma})M_{22}^{-1}(\xi^*,\bar{\gamma})M_{12}^T(\xi^*,\bar{\gamma})|}\right)^{1/s}.}$$

Equivalence theorem

Kiefer and Wolfowitz type (D_S-optimality)

Theorem (Equivalence Theorem: D_S-criterion)

For a localized parameter vector $(\bar{\gamma})$, the following properties are equivalent:

1. ξ^* *is D_s-optimal;*
2. *let us call $A^T = (I_s \, 0)$, then $\forall x \in \mathcal{X}$*

$$\operatorname{tr}\left[M(\xi^*,\bar{\gamma})^{-1}A(A^T M(\xi^*,\bar{\gamma})^{-1}A)^{-1}A^T M(\xi^*,\bar{\gamma})^{-1}m(x,\bar{\gamma})\right] \le s;$$

3. *over all $\xi \in \Xi$, ξ^* minimize*

$$\max_{x\in\mathcal{X}} \operatorname{tr}\left[M(\xi^*,\bar{\gamma})^{-1}A(A^T M(\xi^*,\bar{\gamma})^{-1}A)^{-1}A^T M(\xi^*,\bar{\gamma})^{-1}m(x,\bar{\gamma})\right].$$

JOHANNES KEPLER UNIVERSITY LINZ

Robust D_s- and D_A-optimality

(a pseudo-Bayesian approach)

Objective function:

$$\Psi(\xi;\mathcal{B},A,\gamma) = \int_{\mathcal{B}} \log\det[A^T\{M(\xi,\gamma)\}^{-1}A]^{-1}\,\mathrm{d}F(\gamma), \tag{6}$$

where $\mathcal{B} \subset \mathbb{R}^r$ is the space of possible parameter values and $F(\gamma)$ is a proper prior distribution function for γ.

Definition

For the comparison of designs we define robust D_s-*Efficiency* of the design ξ with respect to the design ξ^* as the ratio

$$\left(\frac{\int_{\mathcal{B}} \log\det[M_{11}(\xi,\gamma) - M_{12}(\xi,\gamma)M_{22}^{-1}(\xi,\gamma)M_{12}^T(\xi,\tilde{\gamma})]\,\mathrm{d}F(\gamma)}{\int_{\mathcal{B}} \log\det[M_{11}(\xi^*,\gamma) - M_{12}(\xi^*,\gamma)M_{22}^{-1}(\xi^*,\gamma)M_{12}^T(\xi^*,\gamma)]\,\mathrm{d}F(\gamma)}\right)^{1/s}.$$

JOHANNES KEPLER UNIVERSITY LINZ

Robust D_s- and D_A-optimality

(a pseudo-Bayesian approach)

Objective function:

$$\Psi(\xi;\mathcal{B},A,\gamma) = \int_{\mathcal{B}} \log\det[A^T\{M(\xi,\gamma)\}^{-1}A]^{-1}\,\mathrm{d}F(\gamma), \qquad (6)$$

where $\mathcal{B} \subset \mathbb{R}^r$ is the space of possible parameter values and $F(\gamma)$ is a proper prior distribution function for γ.

Definition

For the comparison of designs we define robust D_s*-Efficiency* of the design ξ with respect to the design ξ^* as the ratio

$$\left(\frac{\int_{\mathcal{B}} \log\det[M_{11}(\xi,\gamma) - M_{12}(\xi,\gamma)M_{22}^{-1}(\xi,\gamma)M_{12}^T(\xi,\tilde{\gamma})]\,\mathrm{d}F(\gamma)}{\int_{\mathcal{B}} \log\det[M_{11}(\xi^*,\gamma) - M_{12}(\xi^*,\gamma)M_{22}^{-1}(\xi^*,\gamma)M_{12}^T(\xi^*,\gamma)]\,\mathrm{d}F(\gamma)}\right)^{1/s}.$$

Equivalence theorem

Kiefer and Wolfowitz type (robust D_A-optimality)

Theorem (Equivalence Theorem: robust D_A-criterion)

The following properties are equivalent:

1. ξ^* *is robust* D_A*-optimal;*
2. *for every* $\mathbf{x} \in \mathcal{X}$*, the next inequality holds:*

$$\int_{\mathcal{B}} \mathrm{tr}\,[M(\xi^*,\gamma)^{-1}A(A^T M(\xi^*,\gamma)^{-1}A)^{-1}A^T M(\xi^*,\gamma)^{-1}m(\mathbf{x},\gamma)]\,\mathrm{d}F(\gamma) \le s;$$

3. *over all* $\xi \in \Xi$*, the design* ξ^* *minimizes the function*

$$\max_{x\in\mathcal{X}} \int_{\mathcal{B}} \mathrm{tr}\,[M(\xi^*,\gamma)^{-1}A(A^T M(\xi^*,\gamma)^{-1}A)^{-1}A^T M(\xi^*,\gamma)^{-1}m(\mathbf{x},\gamma)]\,\mathrm{d}F(\gamma).$$

JOHANNES KEPLER UNIVERSITY LINZ

A classical example

from Fedorov, 1971

The model

Let us consider two random variables Y_1 and Y_2 such that:

$$E[y_1(x)] = \beta_0 + \beta_1 x + \beta_2 x^2; \;\; E[y_2(x)] = \beta_3 x + \beta_4 x^3 + \beta_5 x^4$$

for each observation x, $0 \leq x \leq 1$. In the classical formulation (Fedorov, 1971), Y_1 and Y_2 are independent with Gaussian margins.

Our investigations

How do different dependence structures affect the optimal designs?

The joint distribution is, then

$$F_{\mathbf{Y}}(y_1, y_2) = C(\Phi(y_1 - \eta_1(x, \beta)), \Phi(y_2 - \eta_2(x, \beta)); \alpha)$$

where C is respectively say the Gumbel, Clayton and Product copula.

A classical example

from Fedorov, 1971

The model

Let us consider two random variables Y_1 and Y_2 such that:

$$E[y_1(x)] = \beta_0 + \beta_1 x + \beta_2 x^2; \;\; E[y_2(x)] = \beta_3 x + \beta_4 x^3 + \beta_5 x^4$$

for each observation x, $0 \leq x \leq 1$. In the classical formulation (Fedorov, 1971), Y_1 and Y_2 are independent with Gaussian margins.

Our investigations

How do different dependence structures affect the optimal designs?

The joint distribution is, then

$$F_{\mathbf{Y}}(y_1, y_2) = C(\Phi(y_1 - \eta_1(x, \beta)), \Phi(y_2 - \eta_2(x, \beta)); \alpha)$$

where C is respectively say the Gumbel, Clayton and Product copula.

JOHANNES KEPLER UNIVERSITY LINZ

A classical example

from Fedorov, 1971

The model

Let us consider two random variables Y_1 and Y_2 such that:

$$E[y_1(x)] = \beta_0 + \beta_1 x + \beta_2 x^2; \;\; E[y_2(x)] = \beta_3 x + \beta_4 x^3 + \beta_5 x^4$$

for each observation x, $0 \leq x \leq 1$. In the classical formulation (Fedorov, 1971), Y_1 and Y_2 are independent with Gaussian margins.

Our investigations

How do different dependence structures affect the optimal designs?

The joint distribution is, then

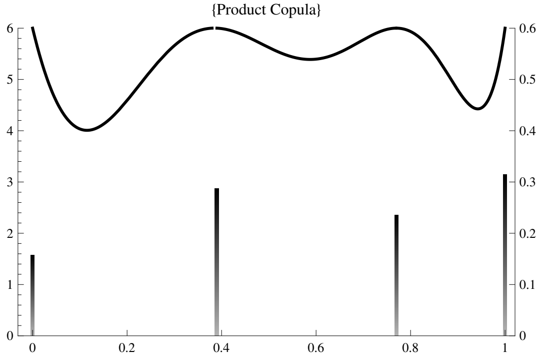
$$F_{\mathbf{Y}}(y_1, y_2) = C(\Phi(y_1 - \eta_1(x,\beta)), \Phi(y_2 - \eta_2(x,\beta)); \alpha)$$

where C is respectively say the Gumbel, Clayton and Product copula.

JOHANNES KEPLER UNIVERSITY LINZ

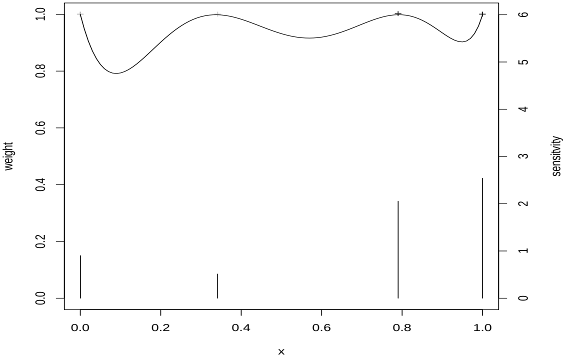
Plots

from Perrone and Müller 2016



Plots

Ds-optimal (Clayton, $\tau = 0.9$) from Perrone, Rappold and Müller 2017



Blocked experiments

of size two

Blocks of size two appear naturally in science, technology and biology.

Examples

- Twin studies, see Valle et al. (2018);
- Body parts (arms, eyes, ears, etc.), see David & Kempton (1996);
- Microarrays, see Bailey (2007).

Material testing

6 blocks of size 2

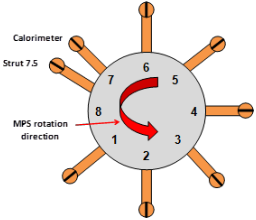


Figure: Arc jet carousel, struts and “wedges” (left) and schematic (right). In addition to the six wedges for holding material samples, the carousel had two further wedges used for temperature measurement.

But first a toy example

motivated by the approach of [Woods and van de Ven(2011)]

Setting

- Poisson regression model with log link, predictor $\beta_0 + \beta_1 x + \beta_2 x^2$;
- Gumbel and Clayton copula, values for Kendall's τ coincide at the levels $\tau_2 = 1/3, 2/3$;
- parameter space $[-1, 1] \times [4, 5] \times [0.5, 1.5]$

OD from an estimating equation approach

$$\xi^\star = \left\{ \begin{array}{ccc} (.03, 1) & (1, .60) & (-.40, .78) \\ .355 & .310 & .335 \end{array} \right\},$$

JOHANNES KEPLER UNIVERSITY LINZ

But first a toy example

motivated by the approach of [Woods and van de Ven(2011)]

Setting

- Poisson regression model with log link, predictor $\beta_0 + \beta_1 x + \beta_2 x^2$;
- Gumbel and Clayton copula, values for Kendall's τ coincide at the levels $\tau_2 = 1/3, 2/3$;
- parameter space $[-1, 1] \times [4, 5] \times [0.5, 1.5]$

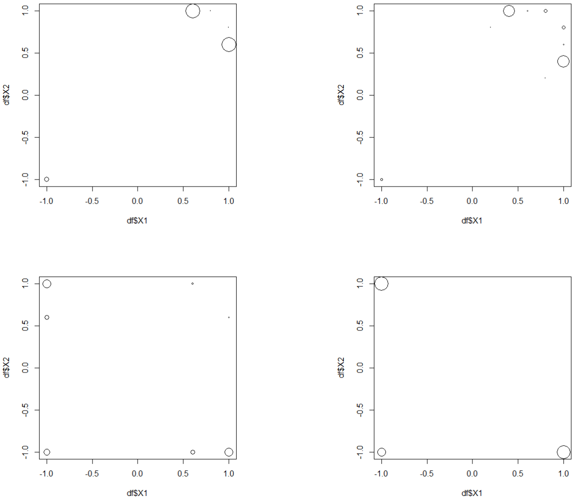
OD from an estimating equation approach

$$\xi^\star = \left\{ \begin{array}{ccc} (.03, 1) & (1, .60) & (-.40, .78) \\ .355 & .310 & .335 \end{array} \right\},$$

JOHANNES KEPLER UNIVERSITY LINZ

Optimal designs for the toy example

rows: Clayton and Gumbel copula; columns levels $\tau_2 = 1/3, 2/3$



Material testing

continued

Setting

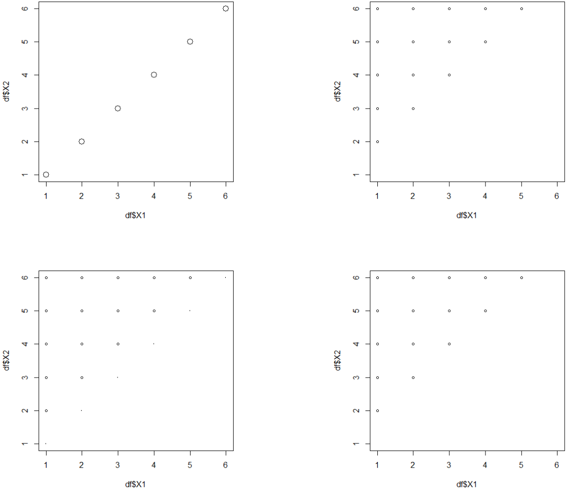
- Marginal model

$$Y_{ij} \sim \text{Bernoulli}(p_{ij}); \quad \log \frac{p_{ij}}{1-p_{ij}} = \beta_0 + \sum_{k=1}^{5} \beta_i x_{ijk} \,,$$

  where Y_{ij} is the binary response from the ith unit in the jth block ($i = 1, 2$; $j = 1, \ldots, n$), p_{ij} is the associated probability of success, x_{ijk} is an indicator variable taking the value 1 if the ith unit in the jth block was assigned treatment k ($k = 1, \ldots, 5$) and 0 otherwise.
- β_0 is the logit for the reference material, with β_k being the difference in expected response, on the logit scale, between the reference material and the kth novel material or treatment.
- Gumbel and Clayton copula, values for Kendall's τ coincide at the levels $\tau_2 = 0, 0.01, 0.1, 0.33$;
- localized solutions (not quasi-Bayesian yet!)

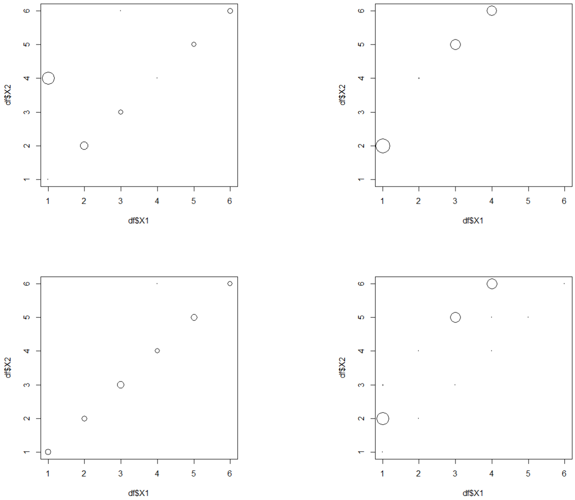
Optimal designs for the real example

rows: Clayton and Gumbel copula; columns levels $\tau_2 = 0, 0.33$; $\beta = \{0,0,0,0,0,0\}$.



Optimal designs for the real example

rows: Clayton and Gumbel copula; columns levels $\tau_2 = 0.01, 0.1$; $\beta = \{0, -1, 2, -3, 4, -5\}$.



JOHANNES KEPLER UNIVERSITY LINZ

Conclusions and next steps

Conclusions

Copulas can be useful tools to construct flexible models in order to:

- separate interblock dependence from marginal behaviour;
- separate asymmetry;
- investigate goodness-of-fit.

Further steps

- Other criteria (particularly for model discrimination);
- Apply other design strategies like multi-stage design procedures;
- Generalize to $m > 2$ with applications in spatial statistics (hydrology, etc.)

JOHANNES KEPLER UNIVERSITY LINZ

Conclusions and next steps

Conclusions

Copulas can be useful tools to construct flexible models in order to:

- separate interblock dependence from marginal behaviour;
- separate asymmetry;
- investigate goodness-of-fit.

Further steps

- Other criteria (particularly for model discrimination);
- Apply other design strategies like multi-stage design procedures;
- Generalize to $m > 2$ with applications in spatial statistics (hydrology, etc.)

JOHANNES KEPLER UNIVERSITY LINZ

Selected references

- Bailey, R. A. (2007), **Designs for two-colour microarray experiments,** *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 56, 365–394.
- V. V. Fedorov. **The design of experiments in the multiresponse case.** Theory of Probability and its Applications, vol **XVI**, (1971), Nr.2.
- M. A. Heise and R. H. Myers. **Optimal Design for Bivariate Logistic Regression.** Biometrics, **52**(2):613–624 (1996).
- E. Perrone and W. G. Müller . **Optimal Designs for Copula Models.**, Statistics **50**(4):917–929 (2016).
- E. Perrone, A. Rappold, W.G. Müller. D_s**-optimality in Copula Models**, Statistical Methods & Applications, online (2016).
- Woods, D. C. and van de Ven, P. (2011), **Blocked Designs for Experiments With Correlated Non-Normal Response,** *Technometrics*, 53, 173–182.

JOHANNES KEPLER UNIVERSITY LINZ

Thanks for the attention!