Bayesian Inference for Conditional Copula models with Continuous and Binary Responses

Radu Craiu

Department of Statistics
University of Toronto

Joint with Avideh Sabeti (Toronto) and Mian Wei (Toronto)

Université Montpellier
April 14, 2014

Outline

Copulas

- Copula functions are used to model dependence between continuous random variables.
- (Sklar,'59) If Y_1, Y_2 are continuous r.v.'s with distribution functions (df) F_1, F_2, there exists an unique copula function $C : [0,1] \times [0,1] \to [0,1]$ such that

$$F_{12}(t,s) = \Pr(Y_1 \leq t, Y_2 \leq s) = C(F_1(t), F_2(s)).$$

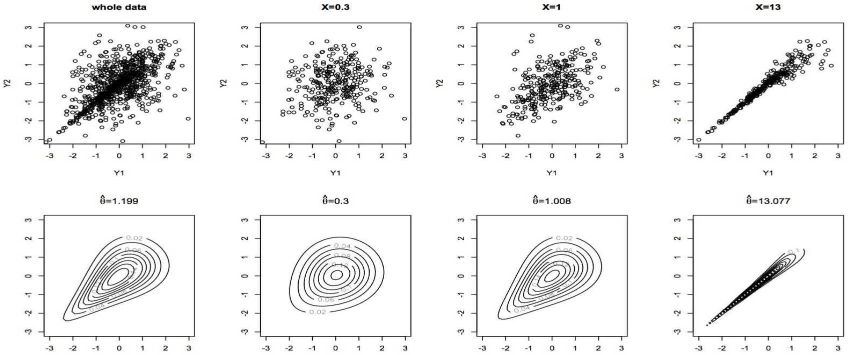
- C is a distribution function on $[0,1]^2$ with uniform margins.
- The copula bridges the marginal distributions with the joint distribution.

Inference for Copula Models

- Early on: abundance of theoretical developments: construction of new copula families and connections with dependence concepts (NA, PQD/NQD, PRD/NRD, etc). Joe ('97), Nelsen ('06).
- Statistical inference for classical parametric copulas
  - Joint Maximum Likelihood: numerical methods
  - Two-stage approach: Joe (JMVA, '05)
  - Semiparametric approach: Genest, Khoudi and Rivest (Bmka, '95)
- Copula selection and goodness-of-fit (Berg, Eur. J. of Finance, '09).
- Emphasis is placed on copula applications in statistics.
- Today: Dynamic regimes of dependence.

Example: Blood pressure

Example It is known that there is a dependence between blood pressure (BP) and body mass index (BMI). What if the dependence varies with subject's age? Can we still use copulas?



Example: Twin Weight

- Twin live births in which both babies survived in the first year of life with mothers of age between 18 and 40.
- Study the dependence between the birth weights of twins, BW_1 and BW_2.
- The gestational age, GA, is an important factor for prenatal growth.
- How does GA influence the dependence between BW_1 and BW_2?

Example: Smoking cessation

- The smoking cessation study of Liu, Daniels and Marcus (JASA '09) :

  Q = smoking cessation (0=No, 1=Yes)
  W = weight change
  X = time spent exercising

- Does exercise weaken the association between smoking status and weight gain?

Conditional Copulas

- The conditional copula of $(Y_1, Y_2)|X = x$, is the conditional joint distribution function of $U = F_{1|X}(Y_1|x)$ and $V = F_{2|X}(Y_2|x)$ given $X = x$ (Patton, Int'l Econ. Rev. '06).
- Consider a random sample $\{x_i, y_{1i}, y_{2i}\}_{1 \le i \le n}$ and suppose $F_{1|X}$ and $F_{2|X}$ are the unknown marginal conditional cdf's.
- The parametric conditional copula model assumes

$$(Y_{1i}, Y_{2i})|X = x_i \sim C(F(Y_{1i}|x_i), F(Y_{2i}|x_i)|\theta(x_i)).$$

- Marginals and copula are conditional on the same variables.

General comments

- Conditional copulas (CC) broaden the range of applications.
  (i) more realistic use of copula models in regression settings.
  (ii) improve the interpretability and understanding of covariate-varying dependence structures.
- Flexible models for the relationship between the association measure (copula parameter, Kendall's tau, etc) and covariate(s) are needed.

A second motivation

- Joint models for high dimensional data.
- The joint distribution of (U_1, U_2, U_3) is modelled using the pair copula model is

$$c(u_1, u_2, u_3) = c_{12}(u_1, u_2)c_{23}(u_2, u_3)c_{13|2}(u_{1|2}, u_{3|2}; u_2)$$

  where $u_{k|2} = Pr(U_k \leq u_k | U_2 = u_2)$.
- Acar, Genest and Neslehova (JMVA, '12) show that wrongly assuming $c_{13|2}$ is the same for all u_2 leads to biased estimators.

Dependence models for CC's in Linear Regression Models

- V_1, V_2 are continuous r.v.'s, $V_i \sim \mathcal{N}(X\beta_i, \sigma_i^2)$, for $i = 1, 2$.
- Jointly,

$$\begin{aligned} f(V_1, V_2|X) &= \prod_{i=1}^{2} \frac{1}{\sigma_i} \phi\left(\frac{V_i - X\beta_i}{\sigma_i}\right) \\ &\times \quad c^{(1,1)} \left\{ \Phi\left(\frac{V_1 - X\beta_1}{\sigma_1}\right), \Phi\left(\frac{V_2 - X\beta_2}{\sigma_2}\right) \middle| \theta(X) \right\}, \end{aligned}$$

  where $c^{(a,b)}(u, v|\theta) = \partial^{a+b} C(u, v|\theta)/\partial u^a \partial v^b$, for all $0 \leq a, b \leq 1$.
- Choose g such that $g(\theta(x_i)) = \eta(x_i)$, where $\eta : \mathbf{R} \to \mathbf{R}$ is the calibration function in inferential focus.

Calibration Model

- A modified version of the cubic spline model of Smith and Kohn (J. Econometrics, '96), Fan et al. (JCGS, '10)

$$\eta(z) = \sum_{j=0}^{3} \alpha_j z^j + \sum_{k=1}^{K} \psi_k (z - \gamma_k)_+^3.$$

- Number and location of knots is influential.

Why a Bayesian approach?

- Joint modelling avoids the propagation of errors.
- Samples from $\pi(\omega|\mathcal{D})$ lead to finite sample variance estimates, pointwise credible regions, computation of model selection criteria.
- Allows data-driven choice of knots location.
- Can use Bayesian model averaging to account for model uncertainty.
- Requires careful examination of the prior's influence.

Bayesian Curve Fitting with Cubic Splines

- Alternative to reversible-jump MCMC approach.
- Partition the range of X into K_{max} intervals, I_k, and introduce auxiliary variables

$$\zeta_k = \begin{cases} 1 & \text{if there is a knot } \gamma_k \text{ in } I_k \text{ and } \psi_k \neq 0 \\ 0 & \text{if there is no knot in } I_k \text{ and } \psi_k = 0 \end{cases}$$

- $\eta(z) = \sum_{j=0}^{3} \alpha_j z^j + \sum_{k=1}^{K_{max}} \psi_k \zeta_k (z - \gamma_k)^3{}_+,$

Bayesian Curve Fitting with Cubic Splines (cont'd)

- Let $|\zeta| = \sum_{k=1}^{K_{max}} \zeta_k$ be the number of knots that are used in the model and set:

  (i) $\lambda \sim \text{Binomial}(K_{max}, p = 0.5)$.

  (ii) $p(|\zeta| \mid \lambda) \propto \frac{\lambda^{|\zeta|}}{|\zeta|!} \mathbf{1}_{\{|\zeta| \leq K_{max}\}}$

  (iii) $p(\zeta \mid |\zeta|) = \binom{K_{max}}{|\zeta|}^{-1}$, all configurations are equally likely.

- $p(\zeta) = p(\zeta \mid |\zeta|)p(|\zeta|)$.

Bayesian Curve Fitting with Cubic Splines (cont'd)

- Prior distributions:

$$
\begin{aligned}
\vec{\alpha} &\sim MVN(0, 10\mathbf{I}_4) \\
\vec{\psi} &\sim N(0, 10\mathbf{I}_{K_{max}}) \\
\gamma_j &\sim \text{Uniform}[I_j],\ j = 1, \dots, K_{max}
\end{aligned}
$$

Prior specification

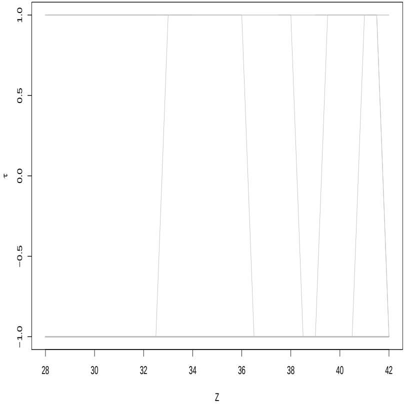


Figure: 500 curves generated from the prior spline model with $K_{max} = 4$.

Prior specification

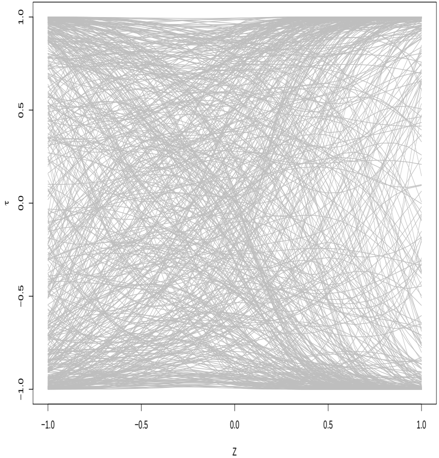


Figure: 500 curves generated from the prior spline model with $K_{max} = 4$. Covariate is standardized.

Multiple covariates

- Numerical problems when the number of covariates is large.
- Additive model approach

$$\eta(z_1, \ldots, z_r) = \sum_{j=1}^{r} \eta_j(z_j)$$

- Additivity is not preserved when changing dependence measure.
- When r is large we may want to use simpler models (piecewise constant).

Intermezzo

- So far:
  - Motivation for CC models
  - Continuous Response Case
  - Calibration function model
  - Multiple covariates.
- Next:
  - Mixed Response Case
  - Model selection issues
  - Simulation results.

Dependence models for CC's in Nonlinear Regression Models

- Measured outcome consists of a binary, Q, and a continuous r.v. W
- Assume a logistic regression model for the binary response.
- For subject j ($j = 1, \ldots, n$) we observe:

$Q_j \sim Bernoulli\left(\frac{\exp(\beta_1 X_j)}{1+\exp(\beta_1 X_j)}\right)$ the binary observed outcome

$W_j \sim N(\beta_2 X_j, \sigma^2)$: the continuous observed outcome

X_j : observed covariate

- Conditional copula dependence model for $Q, W|X$.

Dependence models for CC's in Nonlinear Regression Models (cont'd)

- Estimation for probit/logistic model involves the use of a latent normal/logistic random variable - the binary response variable is equal to the sign of the latent variable.
- Consider $Q_j = \mathbf{1}\{Y_j > 0\}$ where Y_j is a latent logistic random variable with density $f_L(y|A) = \frac{\exp(A-y)}{(1+\exp(A-y))^2}$.
- The latent variable Y is introduced to facilitate computation (for EM, DA).
- Sometimes Y offers a deeper interpretation of the model.
- Marginalization preserves the copula model for $(Y, W)|X$.

The Statistical Model

- The contribution of the jth sample to the observed-data likelihood

$$
\begin{aligned}
\Pr(Q_j = a, W_j | X_j, \omega) &= \frac{\phi\{(W_j - X_j\beta_2)/\sigma_2\}}{\sigma_2} \times \\
\times \quad & \left[c^{(0,1)} \left\{ \frac{\exp(aX_j\beta_1)}{1+\exp(X_j\beta_1)}, \Phi\left(\frac{W_j - X_j\beta_2)}{\sigma_2}\right) \middle| \theta(X_j) \right\} \right]^{1-a} \\
\times \quad & \left[1 - c^{(0,1)} \left\{ \frac{\exp\{(1-a)X_j\beta_1\}}{1+\exp(X_j\beta_1)}, \Phi\left(\frac{W_j - X_j\beta_2}{\sigma_2}\right) \middle| \theta(X_j) \right\} \right]^{a},
\end{aligned}
$$

where $a \in \{0, 1\}$ and ω represents the vector of all the parameters involved in the model.

The Statistical Model

- We assume that the dependence between the latent variable Y and W is characterized by the same conditional copula $C\{\cdot,\cdot|\theta(X)\}$.
- The contribution of the jth sample to the *complete data likelihood* would be (if the Y's were observed)

$$\begin{aligned} f(Y_j, W_j|X_j,\omega) &= f_L\left(Y_j|X_j\beta_1\right)\frac{1}{\sigma_2}\phi\left(\frac{W_j - X_j\beta_2}{\sigma_2}\right) \\ &\times \; c^{(1,1)}\left\{F_L\left(Y_j|X_j\beta_1\right), \Phi\left(\frac{W_j - X_j\beta_2}{\sigma_2}\right)\Bigg|\theta(X_j)\right\} \end{aligned}$$

- $\Pr(Q_j = 0, W_j|X_j,\omega) = \int\limits_{-\infty}^{0} f(Y_j, W_j|X_j,\omega)dY_j$

Remarks

- The usual dependence measures can be used in the bivariate continuous and mixed outcome models.
- If at least one variable is discrete, the dependence parameters are functions of all the parameters in the model,
- Conditional on X, the population Kendall's tau is

$$\begin{aligned}\tau(\omega|X) &= 4\mathrm{E}\{H(Q,W|X)|X\} - 1 = \frac{3 + 2\exp(\beta_1 X) + 3\exp(2\beta_1 X)}{\{1+\exp(\beta_1 X)\}^2} \\ &- 4\int_{\mathbf{R}} \left[\frac{1}{\sigma_2}\phi\left(\frac{w - X\beta_2}{\sigma_2}\right) \times \right. \\ &\times \left. C\left\{\frac{1}{1+\exp(X\beta_1)}, \Phi\left(\frac{w - X\beta_2}{\sigma_2}\right) \middle| \theta(X)\right\}\right] \mathrm{d}w,\end{aligned}$$

where $H(\cdot,\cdot|X)$ is the conditional joint cdf of (Q, W) given X.

Computation Algorithm

- The posterior distribution π is not analytically tractable.
- The sampling scheme requires to alternate between:
  i) sampling the latent variables from their conditional distribution and
  ii) sampling from the conditional posterior distribution of each parameter (given the complete data).

Computation Algorithm (cont'd)

- For most components (Y's, β's, $\log(\sigma)$, α's and ψ's) we use Metropolis-Hastings updates:

Step I Sample a proposal $\tilde{\omega} \sim q(\cdot|\omega_t, s_\omega)$

Step II $\omega_{t+1} = \tilde{\omega}$ with probability $\min\{1, \frac{\pi(\tilde{\omega}|EE)q(\omega_t|\tilde{\omega},s_\omega)}{\pi(\omega_t|EE)q(\tilde{\omega}|\omega_t,s_\omega)}\}$; otherwise $\omega_{t+1} = \omega_t$.

- Choice of s_ω is important and tuning can be time-consuming. We use adaptive MCMC to tune the s_ω's "on the go" .

Two Selection Problems

- Selection of the copula family among a set of candidates.
- Determine whether a parametric calibration function (esp. constant) is suitable.
- Deviance Information Criterion (DIC - Spiegelhalter et al, JRSSB '02).
- Cross-validated marginal likelihood criterion (Geisser and Eddy, JASA '79).

Selection via DIC

- The DIC is defined as

$$DIC(\mathcal{M}) = 2E[\Delta(\omega)|\mathcal{D}_{obs}] - \Delta(E[\omega|\mathcal{D}_{obs}]) \qquad (1)$$

  where the model deviance $\Delta(\omega) = -2\ln p(\mathcal{D}_{obs}|\omega, \mathcal{M})$
- All the required expectations in (1) can be computed using Monte Carlo samples.
- Models with the lowest DIC value are preferred.

A CV Marginal Likelihood Criterion

- The criterion computed under model $\mathcal{M}$ is

$$H(\mathcal{M}) = \sum_{i=1}^{n} \log p(Q_i, W_i|\mathcal{D}_{obs,-i}, \mathcal{M})$$

- $$E[f(Q_i, W_i|\vec{\zeta}, \vec{\omega})^{-1}] = \frac{1}{p(\mathcal{D}_{obs})} \int \frac{f(\mathcal{D}_{obs}|\vec{\zeta}, \vec{\omega})p(\vec{\zeta}, \vec{\omega})}{f(Q_i, W_i|\vec{\zeta}, \vec{\omega})} d\vec{\zeta} d\vec{\omega}$$

$$= \frac{1}{p(\mathcal{D}_{obs})} \int f(\mathcal{D}_{obs,-i}|\vec{\zeta}, \vec{\omega})p(\vec{\zeta}, \vec{\omega}) d\vec{\zeta} d\vec{\omega} = \frac{p(\mathcal{D}_{obs,-i})}{p(\mathcal{D}_{obs})}$$

$$= \frac{1}{p(Q_i, W_i|\mathcal{D}_{obs,-i})}$$

- $p(Q_i, W_i|\mathcal{D}_{obs,-i}, \mathcal{M}) \approx \left[\frac{1}{M} \sum_{m=1}^{M} \frac{1}{f(Q_i, W_i|\vec{\zeta}^{(m)}, \vec{\omega}^{(m)})}\right]^{-1}$,

Simulations

- We performed a large number of simulations to study:
  - the spline model's flexibility in capturing non-linear trends in the calibration function
  - the power to select the correct copula model and to determine the form of the calibration function.
  - the use of additive models for two covariates.
  - variable selection using two criteria.

Simulations

- We have generated data under the Clayton copula family for two sample sizes, $n = 150, 450$. Three dependence patterns are created using the following calibration functions:

$$\begin{aligned} &\text{C1: } \eta(z) = \ln(3); \\ &\text{C2: } \eta(z) = \ln\{0.07z^6 - 0.37(z+1)(z-0.5) + 0.3\}; \\ &\text{C3: } \eta(z) = \ln\{4.5 - 1.5\sin(\pi z)\}. \end{aligned}$$

- The data is analyzed using three copula families: Clayton, Frank and Gumbel with the corresponding link functions $g_C(x) = \ln(1 + x)$, $g_F(x) = x$ and $g_G(x) = \ln(x - 1)$, respectively.

Simulation Results

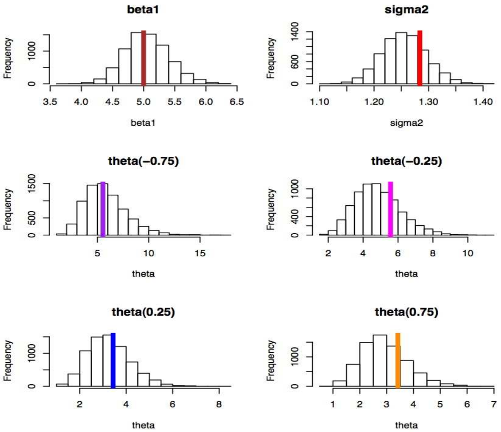
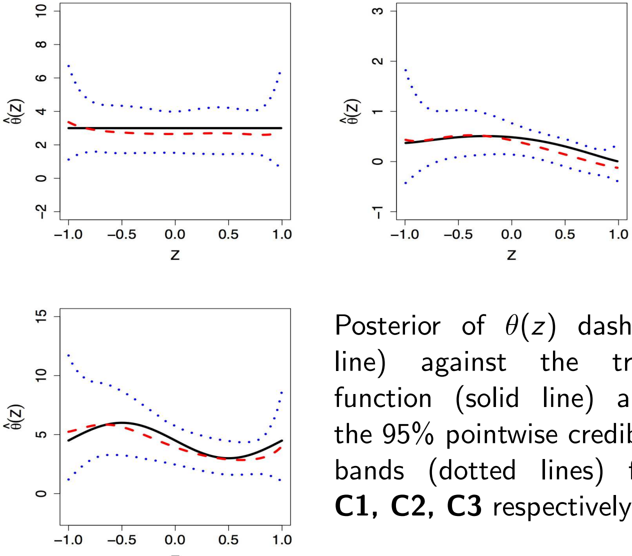


Figure: Histograms of posterior samples obtained from all the parallel MCMC chains when data is generated using a Clayton copula under scenario **C3**. The solid line marks the true parameter value.

Simulation Results



Posterior of $\theta(z)$ dashed line) against the true function (solid line) and the 95% pointwise credible bands (dotted lines) for **C1, C2, C3** respectively.

Estimation Results

Copula	$n = 150$ IBias^2	IVar	IMSE	$n = 450$ IBias^2	IVar	IMSE
	Scenario C1					
Clayton	**0.91**	**4.56**	**5.47**	**0.19**	**1.56**	**1.75**
Frank	23.84	2.90	26.74	28.08	1.94	30.02
Gumbel	5.36	11.62	16.98	4.80	4.18	8.98
	Scenario C2					
Clayton	**0.78**	**6.46**	**7.24**	**0.1**	**2.32**	**3.43**
Frank	4.07	5.94	10.01	3.65	1.98	5.63
Gumbel	18.70	16.64	35.34	17.35	14.61	31.96
	Scenario C3					
Clayton	**0.53**	**4.02**	**4.55**	**0.08**	**1.23**	**1.31**
Frank	21.00	2.25	23.25	22.24	1.43	23.67
Gumbel	2.67	10.88	13.55	5.83	3.43	9.26

Copula Selection

Copula Selection

Percentage of selecting
the Clayton family

$\Downarrow$

Copula Selection

Percentage of selecting the Clayton family

$\Downarrow$

Criteria	n	Frank	Gumbel
Scenario C1			
DIC	150	87%	70%
	450	98%	99%
CVML	150	62%	67%
	450	100%	91%
Scenario C2			
DIC	150	100%	92%
	450	98%	99%
CVML	150	99%	100%
	450	100%	99%
Scenario C3			
DIC	150	94%	80%
	450	100%	98%
CVML	150	86%	60%
	450	100%	97%

Copula Selection

Percentage of selecting
the Clayton family

$\Downarrow$

Criteria	n	Frank	Gumbel
Scenario C1			
DIC	150	87%	70%
	450	98%	99%
CVML	150	62%	67%
	450	100%	91%
Scenario C2			
DIC	150	100%	92%
	450	98%	99%
CVML	150	99%	100%
	450	100%	99%
Scenario C3			
DIC	150	94%	80%
	450	100%	98%
CVML	150	86%	60%
	450	100%	97%

Copula Selection

Calibration Selection

Percentage of selecting the Clayton family

⇓

Criteria	n	Frank	Gumbel
Scenario C1			
DIC	150	87%	70%
	450	98%	99%
CVML	150	62%	67%
	450	100%	91%
Scenario C2			
DIC	150	100%	92%
	450	98%	99%
CVML	150	99%	100%
	450	100%	99%
Scenario C3			
DIC	150	94%	80%
	450	100%	98%
CVML	150	86%	60%
	450	100%	97%

Copula Selection

Percentage of selecting
the Clayton family

$\Downarrow$

Criteria	n	Frank	Gumbel
Scenario C1			
DIC	150	87%	70%
	450	98%	99%
CVML	150	62%	67%
	450	100%	91%
Scenario C2			
DIC	150	100%	92%
	450	98%	99%
CVML	150	99%	100%
	450	100%	99%
Scenario C3			
DIC	150	94%	80%
	450	100%	98%
CVML	150	86%	60%
	450	100%	97%

Calibration Selection

Percentage of selecting
the calibration function

$\Downarrow$

Copula Selection

Percentage of selecting
the Clayton family

⇓

Criteria	n	Frank	Gumbel
Scenario C1			
DIC	150	87%	70%
	450	98%	99%
CVML	150	62%	67%
	450	100%	91%
Scenario C2			
DIC	150	100%	92%
	450	98%	99%
CVML	150	99%	100%
	450	100%	99%
Scenario C3			
DIC	150	94%	80%
	450	100%	98%
CVML	150	86%	60%
	450	100%	97%

Calibration Selection

Percentage of selecting
the calibration function

⇓

Criteria	n	Spline Calibration	Constant Calibration
Scenario C1			
DIC	150	65%	**35%**
	450	49%	**51%**
CVML	150	66%	**34%**
	450	35%	**65%**
Scenario C2			
DIC	150	**93%**	7%
	450	**87%**	13%
CVML	150	**96%**	4%
	450	**100%**	0%
Scenario C3			
DIC	150	**78%**	22%
	450	**75%**	25%
CVML	150	**66%**	34%
	450	**96%**	4%

Copula Selection

Percentage of selecting
the Clayton family

$\Downarrow$

Criteria	n	Frank	Gumbel
Scenario C1			
DIC	150	87%	70%
	450	98%	99%
CVML	150	62%	67%
	450	100%	91%
Scenario C2			
DIC	150	100%	92%
	450	98%	99%
CVML	150	99%	100%
	450	100%	99%
Scenario C3			
DIC	150	94%	80%
	450	100%	98%
CVML	150	86%	60%
	450	100%	97%

Calibration Selection

Percentage of selecting
the calibration function

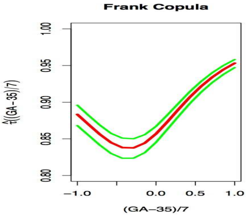
$\Downarrow$

Criteria	n	Spline Calibration	Constant Calibration
Scenario C1			
DIC	150	65%	**35%**
	450	49%	**51%**
CVML	150	66%	**34%**
	450	35%	**65%**
Scenario C2			
DIC	150	**93%**	7%
	450	**87%**	13%
CVML	150	**96%**	4%
	450	**100%**	0%
Scenario C3			
DIC	150	**78%**	22%
	450	**75%**	25%
CVML	150	**66%**	34%
	450	**96%**	4%

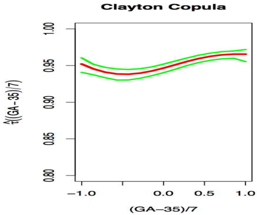
Twin Data Example

- Twin live births in which both babies survived in the first year of life with mothers of age between 18 and 40.
- Study the dependence between the birth weights of twins, BW_1 and BW_2.
- The gestational age, GA, is an important factor for prenatal growth and is therefore chosen as the covariate. We consider a random sample of 30 twin live births for each gestational age (in weeks) between 28 to 45.

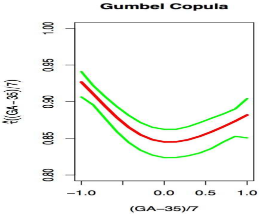
Twin Data Example



DIC=10449



DIC=14810

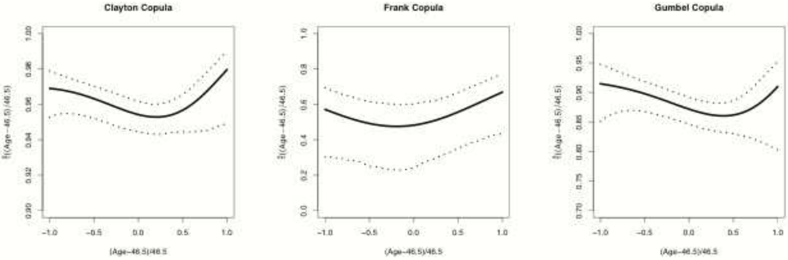


DIC=6.97×10^7

Example: Burn Injury

- $n = 981$ burn injury cases
- Occurrence of death $\Rightarrow Q = 1$ for death and $Q = 0$ for survival
- Total burn area $\Rightarrow W = \log(\text{burn area} + 1)$
- Patient's age
- How age effects the dependence between the severity of burn injury and the probability of death?

Burn Injury



Criteria	n	Clayton	Frank	Gumbel
DIC	981	6865.483	7082.946	6844.854
CVML	981	-3432.229	-3540.972	-3422.06

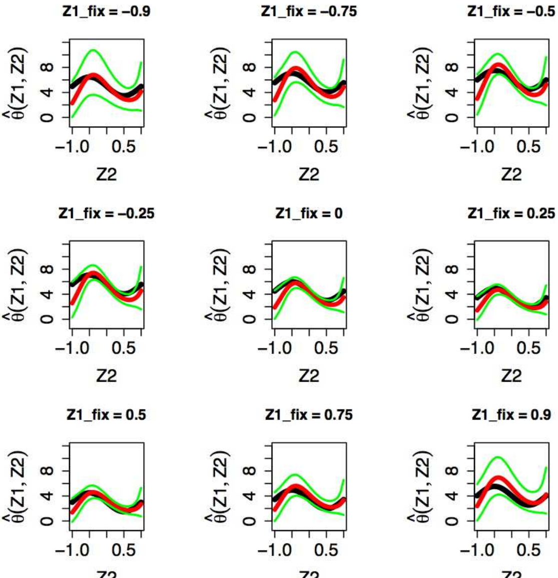
$\Uparrow$

chosen

Additive Model

- Z_1 & Z_2 are iid $N(0, 0.35^2)$.
- $\eta(z_1, z_2) = 4.5 - 1.5\sin(z_1\pi) - 1.5\sin(z_2\pi)$

Additive Model



Variable Selection

Variable Selection

- Data were generated under Clayton family; $n = 450$

Variable Selection

- Data were generated under Clayton family; $n = 450$
- X_1 & $X_2 \sim \mathcal{N}(0, 0.35^2)$

Variable Selection

- Data were generated under Clayton family; $n = 450$
- X_1 & $X_2 \sim \mathcal{N}(0, 0.35^2)$
- **C3** : $\eta(z) = \ln[4.5 - 1.5 \sin(\pi z)]$
- **C4** : $\eta(z_1, z_2) = \ln[4.5 - 1.5 \sin(\pi z_1) - 1.5 \sin(\pi z_2)]$

Variable Selection

- Data were generated under Clayton family; $n = 450$
- X_1 & $X_2 \sim \mathcal{N}(0, 0.35^2)$
- **C3** : $\eta(z) = \ln[4.5 - 1.5\sin(\pi z)]$
- **C4** : $\eta(z_1, z_2) = \ln[4.5 - 1.5\sin(\pi z_1) - 1.5\sin(\pi z_2)]$

- Percentage of selecting the model with correct number of independent variables:

Variable Selection

- Data were generated under Clayton family; $n = 450$
- X_1 & $X_2 \sim \mathcal{N}(0, 0.35^2)$
- **C3** : $\eta(z) = \ln[4.5 - 1.5\sin(\pi z)]$
- **C4** : $\eta(z_1, z_2) = \ln[4.5 - 1.5\sin(\pi z_1) - 1.5\sin(\pi z_2)]$

- Percentage of selecting the model with correct number of independent variables:

Criteria	n	1-Covariate	2-Covariate
Scenario C3			
DIC	450	**96%**	4%
CVML	450	**96%**	4%
Scenario C4			
DIC	450	0%	**100%**
CVML	450	2%	**98%**

Conclusions

- Bayesian analysis for conditional copula models with bivariate responses $\rightarrow$ ... multivariate responses?
- Copula selection via DIC ($\checkmark$) and LP ($\checkmark$).
- Validation of a constant calibration function for continuous ($\checkmark$) and mixed ($\times$) responses.
- Additive models for multivariate predictors are promising.
- CC models offer a gateway into joint modelling of high dimensional data using low-dimensional copulas; the performance of this approach remains untested.

References

- * Avideh Sabeti, Mian Wei and R.C (2014) *Additive models for Bivariate Conditional Copulas.* Preprint.
- ▶ Elif Acar, R.C. and Fang Yao (2014) *Statistical Testing of Covariate Effects in Conditional Copula Models* Electr. J. of Statist.
- * R.C. and Avideh Sabeti (2012). *In mixed company: Bayesian Inference for Bivariate Conditional Copula Models with Discrete and Continuous Outcomes.* JMVA.
- ▶ Elif Acar, R.C. and Fang Yao (2011). *Dependence Calibration in Conditional Copulas: A Nonparametric Approach.*, Biometrics.