Inference on Copulas
When Ignorance is Bliss

Johan Segers\textsuperscript{1}  Christian Genest\textsuperscript{2}

\textsuperscript{1}Université catholique de Louvain  
Institut de statistique

\textsuperscript{2}Université Laval  
Département de mathématique et de statistique

Graybill VIII – 6th EVA (Fort Collins, June 2009)
Puzzle

iid sample \((X_i, Y_i)\) from bivariate normal distribution.

\[
\begin{align*}
E(X) &= \eta + \theta \\
E(Y) &= \eta
\end{align*}
\]

\(\theta\) parameter of interest

\(\eta\) nuisance parameter

Estimating \(\theta = E(X) - E(Y)\) when \(\eta\) is known/unknown:

\[
\hat{\theta}_n = \begin{cases} 
\bar{X}_n - \bar{Y}_n & \text{if } \eta \text{ is unknown} \\
\bar{X}_n - \eta & \text{if } \eta \text{ is known}
\end{cases}
\]

Exercice: If \(\rho > \frac{\sigma_Y}{2\sigma_X}\), then

\[
\text{var}(\bar{X}_n - \bar{Y}_n) < \text{var}(\bar{X}_n - \eta)
\]
Take care when using information on nuisance parameters (2)

Solution
If $\eta = E(Y)$ is known, then the MLE for $\theta = E(X) - E(Y)$ is

$$\hat{\theta}_{n, \text{MLE}}(\eta) = \bar{X}_n - \eta - \frac{\text{cov}(X, Y)}{\text{var}(X)}(\bar{Y}_n - \eta)$$

naive estimator

The MLE makes optimal use of information on $\eta$:

$$\lim_{n \to \infty} \frac{\text{var}(\hat{\theta}_{n, \text{MLE}}(\eta))}{\text{var}(\bar{X}_n - \bar{Y}_n)} \leq 1$$

$\eta$ unknown
For infinite-dimensional parameters, MLE’s may be unknown – then what?

- margins
- (tail) dependence
- “annual maxima”
- EVD
- “domain of attraction”
- GPD
- idem
- our focus
- nuisance
- interest
No method is perfect...

- **Fully parametrically**
  1. Model dependence parametrically
  2. Estimate all parameters jointly by (pseudo)-maximum likelihood

  What if model is wrong?

- **Semi/non-parametrically**: “Plug-in” method
  1. Estimate marginal distributions – *How?*
  2. Use these estimates in order to standardize margins to a convenient distribution
  3. Estimate (tail) dependence in some nonparametric way

  Efficiency?
What we did

- Derive **asymptotic distribution** of plug-in estimators for copulas

  - **Q** Why copulas?
  - **A** Many dependence measures are copula functionals
    (Spectral measure, stable tail dependence function, ...)

- **Compare (co)variance functions** of limiting processes when margins are either:
  - known ("ideal" estimator)
  - estimated parametrically
  - estimated nonparametrically (empirical copula)
Main finding: Ignorance is bliss

Good estimators for the margins do not necessarily yield a good plug-in estimator for the copula.

- Even when margins are known, it may be better to pretend they are not and still estimate them.
- Even when margins are known to belong to certain parametric families, it may be better to ignore this information and still estimate them nonparametrically.
Outline

Plug-in estimators

Asymptotics

Margins known vs. unknown

Margins modelled parametrically

Conclusion
Copula framework as a stylized, general set-up

Bivariate random vector \((X, Y)\) with

\[
\Pr(X \leq x, Y \leq y) = C(F(x), G(y))
\]

dependence: copula \(C\) parameter of interest \((\theta)\)
marginal distributions \(F, G\) nuisance parameters \((\eta)\)

- No restrictions on \(C\) besides smoothness
- Various degrees of information on \(F, G\):
  (a) known
  (b) modelled parametrically
  (c) nothing except for continuity
Estimate the copula by plugging in estimators for the margins

If margins $F$ and $G$ are continuous, then

$$C(u, v) = \Pr(F(X) \leq u, G(Y) \leq v)$$

Plug-in estimator:

$$\hat{C}_n(u, v) = \frac{1}{n} \sum_{i=1}^{n} 1(\hat{F}_n(X_i) \leq u, \hat{G}_n(Y_i) \leq v)$$
Estimators for margins may vary according to information available

<table>
<thead>
<tr>
<th>margins</th>
<th>$\hat{F}_n(x)$</th>
<th>$\hat{G}_n(y)$</th>
<th>$\hat{C}_n(u, v)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>known</td>
<td>$F(x)$</td>
<td>$G(y)$</td>
<td>“ideal”</td>
</tr>
<tr>
<td>parametric</td>
<td>$F(x; \hat{\eta}_n)$</td>
<td>$G(y; \hat{\lambda}_n)$</td>
<td>semiparametric</td>
</tr>
<tr>
<td>nonparametric</td>
<td>$\frac{1}{n} \sum_{i=1}^{n} 1(X_i \leq x)$</td>
<td>$\frac{1}{n} \sum_{i=1}^{n} 1(Y_i \leq y)$</td>
<td>empirical copula</td>
</tr>
</tbody>
</table>

Rüschendorf (1976), Deheuvels (1979)
Outline

Plug-in estimators

Asymptotics

Margins known vs. unknown

Margins modelled parametrically

Conclusion
Influence functions for estimators of margins

Assume that uniformly in $x, y$,

$$
\sqrt{n}(\hat{F}_n(x) - F(x)) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi(X_i; x) + o_p(1)
$$

$$
\sqrt{n}(\hat{G}_n(y) - G(y)) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \zeta(Y_i; y) + o_p(1)
$$

the influence functions $\xi(\cdot; x)$ and $\zeta(\cdot; y)$ being

- mean zero
- finite variance
- $P$-Donsker (as function classes in $x$ and $y$)
The influence functions are straightforward to compute

\[ \sqrt{n}(\hat{F}_n(x) - F(x)) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi(X_i; x) + o_p(1) \]

- Margins known: \( \xi(X; x) = 0 \)
- Nonparametric: \( \xi(X; x) = 1(X \leq x) - F(x) \)
- Parametric: \( \hat{F}_n(x) = F(x; \hat{\eta}_{n,1}, \ldots, \hat{\eta}_{n,p}) \). If

\[ \sqrt{n}(\hat{\eta}_{n,r} - \eta_{r}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi_{r}(X_i) + o_p(1) \]

then by the delta-method,

\[ \xi(X; x) = \sum_{r=1}^{p} \psi_{r}(X) \frac{\partial}{\partial \eta_{r}} F(x; \eta) \]
The plug-in estimator is asymptotically normal

**Theorem**

If \( C \) admits continuous first-order partial derivatives, then in \( \ell_\infty([0, 1]^2) \), denoting \( x = F^{-1}(u) \) and \( y = G^{-1}(v) \),

\[
\sqrt{n}(\hat{C}_n(u, v) - C(u, v)) \xrightarrow{d} \mathcal{G}(1_{(-\infty, x] \times (-\infty, y]}),
\]

\[
\text{C only}
\]

\[
- \dot{C}_1(u, v) \mathcal{G}(\xi(\cdot; x)) - \dot{C}_2(u, v) \mathcal{G}(\zeta(\cdot; y))
\]

\[
\text{estimating } x = F^{-1}(u) \quad \text{estimating } y = G^{-1}(v)
\]

where \( \mathcal{G} \) is a tight, mean-zero Gaussian process with

\[
\text{cov}(\mathcal{G}(f), \mathcal{G}(g)) = \text{cov}(f(X, Y), g(X, Y))
\]
Outline

Plug-in estimators

Asymptotics

Margins known vs. unknown

Margins modelled parametrically

Conclusion
Main theorem provides limit process $C$:

$$\sqrt{n}(\hat{C}_n(u, v) - C(u, v)) \xrightarrow{d} C(u, v)$$

- **Margins known**: “ideal” estimator

$$C_{\text{idl}}(u, v) = \mathbb{G}(1_{(-\infty, F^{-1}(u)] \times (-\infty, G^{-1}(v)]})$$

Distribution depends on $C$ only ($C$-Brownian bridge)

- **Margins completely unknown**: empirical copula

$$C_{\text{emp}}(u, v) = C_{\text{idl}}(u, v) - \dot{C}_1(u, v) C_{\text{idl}}(u, 1) - \dot{C}_2(u, v) C_{\text{idl}}(1, v)$$

What is the net contribution of the two red terms?
The empirical copula may be more efficient than the ideal estimator

**Theorem**

*If* $X$ *is left-tail decreasing (LTD) in* $Y$ *and vice versa then for all* $u, v, s, t \in [0, 1]$

$$\text{cov}(C_{\text{emp}}(u, v), C_{\text{emp}}(s, t)) \leq \text{cov}(C_{\text{idl}}(u, v), C_{\text{idl}}(s, t))$$

**LTD** is a form of positive association:

- monotone regression dependence $\Rightarrow$ LTD $\Rightarrow$ PQD
- Examples:
  - extreme-value copulas (GARRALDA-GUILLEM 2000)
  - positively dependent Gaussian, Frank, Plackett, ...

**Lehmann (1966), Nelson (2006)**
Before you jump to conclusions...

- LTD condition is sufficient but *not necessary*

- For copulas with strong negative association, the conclusion of the theorem may *not* hold
  - Gumbel–Barnett copula

- The result does *not* mean that $C_{\text{emp}}$ is tighter than $C_{\text{idl}}$ in all directions
The variance inequality extends to certain copula-based dependence measures

Plug-in estimator for copula functional $\vartheta = T(C)$:

$$\sqrt{n} \left(\frac{\hat{\vartheta}_n}{T(\hat{C}_n)} - \frac{\vartheta}{T(C)}\right) \xrightarrow{d} T_C(\mathbb{C})$$

**Corollary**

*If $T$ is monotone, then under the conclusion of the theorem,*

$$\text{var}(T_C(C_{emp})) \leq \text{var}(T_C(C_{idl}))$$
Example: Pickands dependence function of an extreme-value copula (1)

Bivariate extreme-value copula

\[ C(u^t, u^{1-t}) = u^{A(t)}, \quad u, t \in (0, 1] \]

**Pickands (1981), Deheuvels (1983)**

- \( A : [0, 1] \rightarrow [0.5, 1] \), convex, \( t \lor (1 - t) \leq A(t) \leq 1 \)
- One-to-one relation between \( A \) and spectral measure, dependence function \( V \), exponent measure \( \mu \), ...
Example: Pickands dependence function of an extreme-value copula (2)

The Pickands dependence function is a functional of the copula:

\[
\frac{1}{A(t)} = \int_0^1 C(u^{1-t}, u^t) \frac{du}{u} \\
- \log A(t) = \int_0^1 (C(u^{1-t}, u^t) - 1(u > e^{-1})) \frac{du}{u \log u} + \Gamma'(1)
\]

Plug-in estimators \( \hat{A}_n(t) = T(\hat{C}_n) \) yield estimators of

- Pickands (1981)
- Capéraà, Fougères & Genest (1997)

**Corollary**

The rank-based Pickands and CFG estimators are more efficient than the “ideal” ones.
Outline

Plug-in estimators

Asymptotics

Margins known vs. unknown

Margins modelled parametrically

Conclusion
Parametric estimators for the margins yield a semiparametric plug-in estimator for the copula

Plugging in estimates $\hat{\eta}_n$ and $\hat{\lambda}_n$ yields

$$\hat{C}_{n,\text{par}}(u, v) = \frac{1}{n} \sum_{i=1}^{n} 1(F(X_i; \hat{\eta}_n) \leq u, G(Y_i; \hat{\lambda}_n) \leq v)$$

Asymptotics depend on:

- the copula
- the parametric models for the margins
- the parameter estimates
Recall: The plug-in estimator is asymptotically normal

**Corollary**
In $\ell_\infty([0,1]^2)$, denoting $x = F^{-1}(u)$ and $y = G^{-1}(v)$,

$$
\sqrt{n} (\hat{C}_{n,\text{par}}(u, v) - C(u, v)) 
\xrightarrow{d} \mathbb{G}(1_{(-\infty,x] \times (-\infty,y]}) 
- \hat{C}_1(u, v) \mathbb{G}(\xi(\cdot;x)) - \hat{C}_2(u, v) \mathbb{G}(\zeta(\cdot;y))
$$

with

- $\mathbb{G}$ a $P$-Brownian bridge
- the influence functions $\xi(\cdot;x)$ and $\zeta(\cdot;y)$ depending on
  - the parametric models for the margins
  - the asymptotic influence functions of $\hat{\eta}_n$ and $\hat{\lambda}_n$
In case of independence, asymptotic variances are computable

**Corollary**
If $C(u, v) = uv$ (independence), then

- *estimator asymptotic variance* $\text{var}(\bar{C}(u, v))$

  - "ideal" $uv(1 - uv)$
  - ranks $uv(1 - uv) - v^2 u(1 - u) - u^2 v(1 - v)$
  - parametric $uv(1 - uv) - v^2 \alpha(u) - u^2 \beta(v)$

with

\[
\alpha(u) = 2 \ E[\xi(X; x) \mathbf{1}(X \leq x)] - E[\xi^2(X; x)], \quad x = F^{-1}(u) \\
\beta(v) = 2 \ E[\zeta(Y; y) \mathbf{1}(Y \leq y)] - E[\zeta^2(Y; y)], \quad y = G^{-1}(v)
\]
In many examples, the empirical copula is more efficient than the semiparametric estimator.

**Corollary**

If $\alpha(u) < u(1 - u)$, then, at independence, the empirical copula beats the semiparametric estimator.

<table>
<thead>
<tr>
<th>Model</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha(u)$ when $\hat{\eta}_n$ is MLE</td>
<td>$\alpha(u)$</td>
</tr>
<tr>
<td>Exponential scale</td>
<td>$(1 - u)^2 \log^2(1 - u)$</td>
</tr>
<tr>
<td>Gumbel location</td>
<td>$u^2 \log^2(u)$</td>
</tr>
<tr>
<td>Normal location</td>
<td>$\varphi^2(\Phi^{-1}(u))$</td>
</tr>
<tr>
<td>Normal location-scale</td>
<td>$\varphi^2(\Phi^{-1}(u)) \left( 1 + \frac{1}{2} (\Phi^{-1}(u))^2 \right)$</td>
</tr>
</tbody>
</table>
Outline

Plug-in estimators

Asymptotics

Margins known vs. unknown

Margins modelled parametrically

Conclusion
Naive usage of knowledge on margins may result in inefficient copula estimators

Good estimators for the margins do not necessarily yield a good plug-in estimator for the copula

- Empirical copula often performs best—but not always.
- Plug-in method is not efficient.

Open problem
In case margins are modelled parametrically, find truly efficient estimators for the dependence structure.

www.uclouvain.be/stat