Inference on Copulas

When Ignorance is Bliss

Johan Segers[1] Christian Genest[2]

[1]Université catholique de Louvain
Institut de statistique

[2]Université Laval
Déptarement de mathématique et de statistique

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

Take care when using information on nuisance parameters (1)

Puzzle

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

$$\mathrm{E}(X) = \eta + \theta \qquad \theta \text{ parameter of interest}$$
$$\mathrm{E}(Y) = \eta \qquad \eta \text{ nuisance parameter}$$

Estimating $\theta = \mathrm{E}(X) - \mathrm{E}(Y)$ when η is *known/unknown*:

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

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

$$\mathrm{var}\big(\bar{X}_n - \bar{Y}_n\big) < \mathrm{var}\big(\bar{X}_n - \eta\big)$$

Take care when using information on nuisance parameters (2)

Solution

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

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

The MLE makes optimal use of information on η:

$$\lim_{n\to\infty} \mathrm{var}\big(\hat{\theta}_{n,\mathrm{MLE}}(\eta)\big) \big/ \underbrace{\mathrm{var}\big(\bar{X}_n - \bar{Y}_n\big)}_{\eta \text{ unknown}} \leqslant 1$$

For infinite-dimensional parameters, MLE's may be unknown – then what?

	margins	*(tail) dependence*
"annual maxima"	EVD	spectral measure, . . .
"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

Outline

Copula framework as a stylized, general set-up

Bivariate random vector (X, Y) with

$$\Pr(X \leqslant x,\ Y \leqslant y) = C\big(F(x), G(y)\big)$$

dependence: copula	C	parameter of interest	(θ)
marginal distributions	F, G	nuisance parameters	(η)

- 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

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

Plug-in estimator:

$$\hat{C}_n(u,v) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\big(\hat{F}_n(X_i) \leqslant u,\, \hat{G}_n(Y_i) \leqslant v\big)$$

Estimators for margins may vary according to information available

margins	$\hat{F}_n(x)$	$\hat{G}_n(y)$	$\hat{C}_n(u,v)$
known	$F(x)$	$G(y)$	"ideal"
parametric	$F(x;\hat{\eta}_n)$	$G(y;\hat{\lambda}_n)$	semiparametric
nonparametric	$\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(X_i \leqslant x)$	$\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(Y_i \leqslant y)$	empirical copula

RÜSCHENDORF (1976), DEHEUVELS (1979)

Outline

Influence functions for estimators of margins

Assume that uniformly in x, y,

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

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

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

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

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

$$\sqrt{n}\big(\hat{\eta}_{n,r} - \eta_r\big) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi_r(X_i) + o_p(1)$$

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 $y = G^{-1}(v)$,

$$\sqrt{n}\big(\hat{C}_n(u,v) - C(u,v)\big) \xrightarrow{d} \underbrace{\mathbb{G}\big(\mathbf{1}_{(-\infty,x]\times(-\infty,y]}\big)}_{C\ \text{only}} - \underbrace{\dot{C}_1(u,v)\,\mathbb{G}\big(\xi(\,\cdot\,;x)\big)}_{\text{estimating } x = F^{-1}(u)} - \underbrace{\dot{C}_2(u,v)\,\mathbb{G}\big(\zeta(\,\cdot\,;y)\big)}_{\text{estimating } y = G^{-1}(v)}$$

where $\mathbb{G}$ is a tight, mean-zero Gaussian process with

$$\mathrm{cov}\big(\mathbb{G}(f),\, \mathbb{G}(g)\big) = \mathrm{cov}\big(f(X,Y),\, g(X,Y)\big)$$

Outline

Ideal estimator vs. empirical copula

Main theorem provides limit process $\mathbb{C}$:

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

- Margins *known*: "ideal" estimator

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

  Distribution depends on C only (C-Brownian bridge)

- Margins *completely unknown*: empirical copula

$$\begin{aligned}&\mathbb{C}_{\text{emp}}(u,v)\\ &\quad = \mathbb{C}_{\text{idl}}(u,v) - \dot{C}_1(u,v)\,\mathbb{C}_{\text{idl}}(u,1) - \dot{C}_2(u,v)\,\mathbb{C}_{\text{idl}}(1,v)\end{aligned}$$

  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 $u, v, s, t \in [0, 1]$

$$\operatorname{cov}\big(\mathbb{C}_{\text{emp}}(u,v),\, \mathbb{C}_{\text{emp}}(s,t)\big) \leqslant \operatorname{cov}\big(\mathbb{C}_{\text{idl}}(u,v),\, \mathbb{C}_{\text{idl}}(s,t)\big)$$

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), NELSEN (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 $\mathbb{C}_{\text{emp}}$ is tighter than $\mathbb{C}_{\text{idl}}$ in all directions

The variance inequality extends to certain copula-based dependence measures

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

$$\sqrt{n}\big(\underbrace{\hat{\vartheta}_n}_{T(\hat{C}_n)} - \underbrace{\vartheta}_{T(C)}\big) \xrightarrow{d} \dot{T}_C(\mathbb{C})$$

Corollary

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

$$\mathrm{var}\big(\dot{T}_C(\mathbb{C}_{emp})\big) \leqslant \mathrm{var}\big(\dot{T}_C(\mathbb{C}_{idl})\big)$$

Example: Pickands dependence function of an extreme-value copula (1)

Bivariate extreme-value copula

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

PICKANDS (1981), DEHEUVELS (1983)

- $A : [0, 1] \to [0.5, 1]$, convex, $t \vee (1 - t) \leqslant A(t) \leqslant 1$
- One-to-one relation between A and spectral measure, dependence function V, exponent measure μ, ...

Example: Pickands dependence function of an extreme-value copula (2)

The Pickands dependence function is a functional of the copula:

$$\frac{1}{A(t)} = \int_0^1 C(u^{1-t}, u^t)\,\frac{\mathrm{d}u}{u}$$

$$-\log A(t) = \int_0^1 \big(C(u^{1-t}, u^t) - \mathbf{1}(u > e^{-1})\big)\,\frac{\mathrm{d}u}{u \log u} + \Gamma'(1)$$

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

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

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

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 $y = G^{-1}(v)$,

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

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* $\mathrm{var}\big(\mathbb{C}(u,v)\big)$
"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

$$\alpha(u) = 2\ \mathrm{E}[\xi(X;x)\,\mathbf{1}(X \leqslant x)] - \mathrm{E}[\xi^2(X;x)], \quad x = F^{-1}(u)$$
$$\beta(v) = 2\ \mathrm{E}[\zeta(Y;y)\,\mathbf{1}(Y \leqslant y)] - \mathrm{E}[\zeta^2(Y;y)], \quad y = G^{-1}(v)$$

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.

model	$\alpha(u)$ when $\hat{\eta}_n$ is MLE
exponential scale	$(1-u)^2 \log^2(1-u)$
Gumbel location	$u^2 \log^2(u)$
normal location	$\varphi^2\big(\Phi^{-1}(u)\big)$
normal location-scale	$\varphi^2\big(\Phi^{-1}(u)\big)\left(1+\frac{1}{2}\big(\Phi^{-1}(u)\big)^2\right)$

Outline

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.