A Simulation Study of Pseudo-Likelihood Information Criteria for Copula Model Selection

Aibat Kossumov
Department of Probability and Mathematical Statistics,
Faculty of Mathematics and Physics,
Charles University

1 Abstract

One of the fundamental problems in dependence modeling is the selection of an appropriate parametric copula model. In [1], it was shown that using the Akaike Information Criterion (AIC) based on the pseudo-log-likelihood is not justified for selecting parametric copula models. As a possible alternative, the authors proposed the information criterion xv_1, based on leave-one-out cross-validation, along with its approximation $\mathrm{xv}_{\mathrm{CIC}}$. In [2], the AIC and $\mathrm{xv}_{\mathrm{CIC}}$ were compared, and only minor differences were observed. In the context of linear model selection, Jun Shao [3] demonstrated that the optimal selection procedure is leave-n_v-out cross-validation, where n_v is of the same order as the sample size n, i.e., $n_v/n \underset{n\to\infty}{\longrightarrow} 1$. This idea is adapted to the context of copula model selection. Its performance is compared with that of AIC, xv_1 and $\mathrm{xv}_{\mathrm{CIC}}$.

2 Used Information Criteria

In this simulation study, we compare four different copula selection methods:

- the Akaike Information Criterion (AIC),
- leave-one-out cross-validation xv_1 and its approximation $\mathrm{xv}_{\mathrm{CIC}}$,
- leave-n_v-out cross-validation xv_{n_v}, where the validation set size n_v is of the same asymptotic order as the total sample size n.

We restrict our attention to the two-dimensional case and copula families with a one-dimensional dependence parameter θ, such as Clayton, Gumbel, Joe, Frank, and Gaussian. Denote by $\mathcal{X}_n = \{\boldsymbol{x}_i\}_{i=1}^n$ a random sample from the joint cdf

$$H(x_1, x_2) = C(F_1(x_1), F_2(x_2)),$$

where C is the copula, and F_1 and F_2 are continuous but unknown marginal cdfs. Also, define

$$\widetilde{\boldsymbol{F}}_n(x_1, x_2) = \left(\widetilde{F}_{n,1}(x_1), \widetilde{F}_{n,2}(x_2)\right),$$

where $\widetilde{F}_{n,k}$ is the $\frac{n}{n+1}$-rescaled empirical cdf of the kth marginal, for $k = 1, 2$. The corresponding pseudo-observations are denoted by ${}^p\mathcal{X}_n = \{{}^p\boldsymbol{x}_i\}_{i=1}^n$, where ${}^p\boldsymbol{x}_i = \widetilde{\boldsymbol{F}}_n(\boldsymbol{x}_i)$.

Note that [4, page 59] discusses why it is sufficient to simulate data from a copula model rather than a full bivariate model.

2.1 Akaike Information Criterion (AIC)

The AIC in the case of a one-dimensional parameter θ is given by:

$$\mathrm{AIC} = 2 \cdot {}^{p}\ell_n(\widehat{\theta}_n) - 2,$$

where ${}^{p}\ell_n$ is the pseudo-log-likelihood, which implicitly depends on the pseudo-observations ${}^{p}\mathcal{X}_n$, and is given by:

$${}^{p}\ell_n(\theta) = \sum_{i=1}^{n} \log[c_\theta({}^{p}\boldsymbol{x}_i)],$$

and $\widehat{\theta}_n = \underset{\theta\in\Theta}{\operatorname{argmax}}\ {}^{p}\ell_n(\theta)$ is the maximum pseudo-likelihood estimator.

2.2 Information Criterion Based on Leave-One-Out Cross-Validation

The selection procedure is based on the following quantity:

$$\mathrm{xv}_1 = \frac{1}{n}\sum_{i=1}^{n} \log\left[c_\theta\left(\widetilde{\boldsymbol{F}}_{(-i)}(\boldsymbol{x}_i)\right)\right]_{\theta=\widehat{\theta}_{(-i)}}, \text{ where} \tag{1}$$

- $\widehat{\theta}_{(-i)} = \underset{\theta\in\Theta}{\operatorname{argmax}}\ \sum_{j\neq i} \log\left[c_\theta\left(\widetilde{\boldsymbol{F}}_{(-i)}(\boldsymbol{x}_j)\right)\right]$,
- $\widetilde{\boldsymbol{F}}_{(-i)}(x_1, x_2) = \left(\widetilde{F}_{(-i),1}(x_1), \widetilde{F}_{(-i),2}(x_2)\right)$, where $\widetilde{F}_{(-i),k}$ is the $\frac{n-1}{n}$-rescaled empirical cdf of the kth marginal, computed from the sample $\mathcal{X}_n$ excluding $\boldsymbol{x}_i$, for $k = 1, 2$.

Since computing (1) is computationally expensive, the authors of [1] recommend using $\mathrm{xv}_{\mathrm{CIC}}$, which is an asymptotically equivalent version and is given by:

$$\mathrm{xv}_{\mathrm{CIC}} = 2 \cdot \left({}^{p}\ell_n(\widehat{\theta}_n) - \widehat{p}_n - \widehat{q}_n - \widehat{r}_n\right), \text{ where} \tag{2}$$

- $\widehat{p}_n = \frac{1}{n\cdot\widehat{J}} \sum_{i=1}^{n} \left[\phi_\theta({}^{p}\boldsymbol{x}_i)\right]^2_{\theta=\widehat{\theta}_n}$,
- $\widehat{q}_n = \frac{1}{n\cdot\widehat{J}} \sum_{i=1}^{n} \left[\phi_\theta({}^{p}\boldsymbol{x}_i)\cdot \widehat{z}_\theta({}^{p}\boldsymbol{x}_i)\right]_{\theta=\widehat{\theta}_n}$,
- $\widehat{r}_n = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\partial \log c_\theta({}^{p}\boldsymbol{x}_i)}{\partial u_1}\cdot(1 - {}^{p}x_{i,1}) + \frac{\partial \log c_\theta({}^{p}\boldsymbol{x}_i)}{\partial u_2}\cdot(1 - {}^{p}x_{i,2})\right]_{\theta=\widehat{\theta}_n}$,
- $\phi_\theta(\boldsymbol{u}) = \frac{\partial \log c_\theta(\boldsymbol{u})}{\partial\theta}$,
- $\widehat{z}_\theta(\boldsymbol{x}) = \frac{1}{n}\sum_{k=1}^{2}\sum_{i=1}^{n} \frac{\partial\phi_\theta({}^{p}\boldsymbol{x}_i)}{\partial u_k}\cdot(\mathbf{1}\{x_k \leq {}^{p}x_{i,k}\} - {}^{p}x_{i,k})$,
- $\widehat{J} = -\frac{1}{n}\sum_{i=1}^{n}\left[\frac{\partial^2 \log c_\theta({}^{p}\boldsymbol{x}_i)}{\partial\theta^2}\right]_{\theta=\widehat{\theta}_n}$.

The generalization of formula (2) to higher dimensions can be found in [4, page 55].

2.3 Information Criterion Based on Leave-n_v-Out Cross-Validation

Inspired by [3], we randomly draw, without replacement, a collection $\mathcal{T}_n$ of $b_n = O(n)$ subsets of $\{1, \ldots, n\}$, each of size n_v, such that $n_v/n \underset{n\to\infty}{\longrightarrow} 1$. Here, the n_v observations are used for validation, while the remaining $n_c = n - n_v$ observations are used for parameter estimation. Denote by $s_v \in \mathcal{T}_n$ the set of indices corresponding to the n_v validation observations. Then define the following quantity:

$$\mathrm{xv}_{n_v} = \frac{1}{n_v b_n} \sum_{s_v \in \mathcal{T}_n} \sum_{i \in s_v} \log \left[c_\theta \left(\widetilde{\boldsymbol{F}}_{(-s_v)}(\boldsymbol{x}_i) \right) \right]_{\theta = \widehat{\theta}_{(-s_v)}}, \text{ where}$$

- $\widehat{\theta}_{(-s_v)} = \underset{\theta\in\Theta}{\operatorname{argmax}} \ \sum_{j \notin s_v} \log \left[c_\theta \left(\widetilde{\boldsymbol{F}}_{(-s_v)}(\boldsymbol{x}_j) \right) \right]$,
- $\widetilde{\boldsymbol{F}}_{(-s_v)}(x_1, x_2) = \left(\widetilde{F}_{(-s_v),1}(x_1), \widetilde{F}_{(-s_v),2}(x_2) \right)$, where $\widetilde{F}_{(-s_v),k}$ is the $\frac{n_c}{n_c+1}$-rescaled empirical cdf of the kth marginal, computed from the sample $\mathcal{X}_n$ excluding $\{\boldsymbol{x}_i : i \in s_v\}$, for $k = 1, 2$.

3 Setup of the Simulation Study

In this simulation study, the following settings were considered:

- The copulas C were chosen from one-dimensional parametric families (Clayton, Gumbel, Joe, Frank, Gaussian).
- Each copula was parameterized using different values of Kendall's tau. Specifically, for $\tau \in \{0.25, 0.5, 0.75\}$, we considered sample sizes $n \in \{100, 250, 500\}$. We also considered cases with weak dependence, $\tau \in \{0.05, 0.1, 0.15, 0.2\}$, and smaller sample sizes, $n \in \{100, 200\}$.
- For each pair of τ and n, we conducted 1000 replications.
- For the calculation of xv_{n_v}, we used $b_n = \lfloor 0.8n \rfloor$ and $n_c = n^{0.9}$.

References

[1] S. Grønneberg, N. L. Hjort. *The copula information criteria*. Scand. J. Stat. **41** (2014) 436–459.

[2] L. A. Jordanger, D. Tjøstheim. *Model selection of copulas: AIC versus a cross validation copula information criterion*. Statist. Probab. Lett. **92** (2014) 249–255.

[3] J. Shao. *Linear model selection by cross-validation*. J. Amer. Statist. Assoc. **88** (1993) 486–494.

[4] L. A. Jordanger *Semiparametric model selection for copulas*. Master's Thesis in Statistics (2013).

4 Model Selection Counts

In the tables of this section, the first column, denoted as d.cop, indicates the true copula from which the data were simulated. Each row corresponds to one of the four selection methods, and the numbers in the cells represent how many times a specific copula (from the columns) was selected by that method across 1000 replications.

4.1 $\tau \in \{0.25, 0.5, 0.75\}$ and $n \in \{100, 250, 500\}$

Here, in most of the 1000 replications, the individual methods were able to select the correct copula. The only case of frequent incorrect model selection occurs in Table 1, where the data were simulated from the Gumbel copula, but $\mathrm{xv}_{\mathrm{CIC}}$ more often selected the Joe copula.

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	860	12	1	46	81
Clayton	xv_1	860	12	1	46	81
Clayton	$\mathrm{xv}_{\mathrm{CIC}}$	807	13	1	58	121
Clayton	xv_{n_v}	851	14	1	52	82
Gumbel	AIC	30	376	362	92	140
Gumbel	xv_1	30	376	362	92	140
Gumbel	$\mathrm{xv}_{\mathrm{CIC}}$	23	336	432	91	118
Gumbel	xv_{n_v}	30	396	346	104	124
Joe	AIC	1	177	777	22	23
Joe	xv_1	1	177	777	22	23
Joe	$\mathrm{xv}_{\mathrm{CIC}}$	1	134	823	24	18
Joe	xv_{n_v}	0	203	753	21	23
Frank	AIC	119	130	38	509	204
Frank	xv_1	119	130	38	508	205
Frank	$\mathrm{xv}_{\mathrm{CIC}}$	82	124	62	534	198
Frank	xv_{n_v}	124	148	33	509	186
Gaussian	AIC	158	189	58	198	397
Gaussian	xv_1	159	189	58	196	398
Gaussian	$\mathrm{xv}_{\mathrm{CIC}}$	109	197	86	218	390
Gaussian	xv_{n_v}	166	210	57	207	360

Table 1: Copula selection using different information criteria ($n = 100$, $\tau = 0.25$)

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	967	0	0	12	21
Clayton	xv_1	967	0	0	12	21
Clayton	xv_{CIC}	928	0	0	24	48
Clayton	xv_{n_v}	970	0	0	13	17
Gumbel	AIC	2	675	192	36	95
Gumbel	xv_1	2	675	192	36	95
Gumbel	xv_{CIC}	0	682	216	41	61
Gumbel	xv_{n_v}	3	688	185	39	85
Joe	AIC	0	154	840	4	2
Joe	xv_1	0	154	840	4	2
Joe	xv_{CIC}	0	118	877	5	0
Joe	xv_{n_v}	0	163	832	5	0
Frank	AIC	17	60	4	742	177
Frank	xv_1	17	60	4	742	177
Frank	xv_{CIC}	5	74	5	783	133
Frank	xv_{n_v}	19	69	2	742	168
Gaussian	AIC	36	157	2	90	715
Gaussian	xv_1	36	157	2	90	715
Gaussian	xv_{CIC}	19	212	3	115	651
Gaussian	xv_{n_v}	36	180	1	111	672

Table 2: Copula selection using different information criteria ($n = 100$, $\tau = 0.50$)

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	990	0	0	5	5
Clayton	xv_1	990	0	0	5	5
Clayton	xv_{CIC}	973	0	0	12	15
Clayton	xv_{n_v}	989	0	0	5	6
Gumbel	AIC	0	806	53	19	122
Gumbel	xv_1	0	806	53	19	122
Gumbel	xv_{CIC}	0	824	72	28	76
Gumbel	xv_{n_v}	0	822	50	18	110
Joe	AIC	0	84	912	4	0
Joe	xv_1	0	84	912	4	0
Joe	xv_{CIC}	0	62	934	4	0
Joe	xv_{n_v}	0	94	902	4	0
Frank	AIC	3	18	0	895	84
Frank	xv_1	3	18	0	894	85
Frank	xv_{CIC}	1	26	0	927	46
Frank	xv_{n_v}	3	22	0	899	76
Gaussian	AIC	8	126	0	45	821
Gaussian	xv_1	8	126	0	45	821
Gaussian	xv_{CIC}	5	178	0	70	747
Gaussian	xv_{n_v}	10	138	0	52	800

Table 3: Copula selection using different information criteria ($n = 100$, $\tau = 0.75$)

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	971	0	0	7	22
Clayton	xv_1	971	0	0	7	22
Clayton	xv_{CIC}	953	0	0	12	35
Clayton	xv_{n_v}	973	0	0	6	21
Gumbel	AIC	2	629	239	44	86
Gumbel	xv_1	2	629	239	44	86
Gumbel	xv_{CIC}	2	604	281	46	67
Gumbel	xv_{n_v}	2	647	227	45	79
Joe	AIC	0	151	847	1	1
Joe	xv_1	0	151	847	1	1
Joe	xv_{CIC}	0	116	882	1	1
Joe	xv_{n_v}	0	163	835	1	1
Frank	AIC	37	44	1	733	185
Frank	xv_1	37	44	1	733	185
Frank	xv_{CIC}	23	52	2	761	162
Frank	xv_{n_v}	35	51	2	743	169
Gaussian	AIC	71	125	1	146	657
Gaussian	xv_1	72	125	1	146	656
Gaussian	xv_{CIC}	52	144	1	157	646
Gaussian	xv_{n_v}	75	133	1	154	637

Table 4: Copula selection using different information criteria ($n = 250$, $\tau = 0.25$)

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	998	0	0	2	0
Clayton	xv_1	998	0	0	2	0
Clayton	xv_{CIC}	997	0	0	2	1
Clayton	xv_{n_v}	998	0	0	2	0
Gumbel	AIC	0	895	63	8	34
Gumbel	xv_1	0	895	63	8	34
Gumbel	xv_{CIC}	0	897	75	8	20
Gumbel	xv_{n_v}	0	902	61	10	27
Joe	AIC	0	35	965	0	0
Joe	xv_1	0	35	965	0	0
Joe	xv_{CIC}	0	27	973	0	0
Joe	xv_{n_v}	0	37	963	0	0
Frank	AIC	1	8	0	940	51
Frank	xv_1	1	8	0	940	51
Frank	xv_{CIC}	0	8	0	952	40
Frank	xv_{n_v}	1	8	0	952	39
Gaussian	AIC	3	68	0	26	903
Gaussian	xv_1	3	68	0	26	903
Gaussian	xv_{CIC}	1	91	0	29	879
Gaussian	xv_{n_v}	3	74	0	27	896

Table 5: Copula selection using different information criteria ($n = 250$, $\tau = 0.50$)

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	1000	0	0	0	0
Clayton	xv_1	1000	0	0	0	0
Clayton	xv_{CIC}	1000	0	0	0	0
Clayton	xv_{n_v}	1000	0	0	0	0
Gumbel	AIC	0	969	5	0	26
Gumbel	xv_1	0	969	5	0	26
Gumbel	xv_{CIC}	0	978	7	1	14
Gumbel	xv_{n_v}	0	970	5	0	25
Joe	AIC	0	4	996	0	0
Joe	xv_1	0	4	996	0	0
Joe	xv_{CIC}	0	4	996	0	0
Joe	xv_{n_v}	0	4	996	0	0
Frank	AIC	0	0	0	999	1
Frank	xv_1	0	0	0	999	1
Frank	xv_{CIC}	0	0	0	999	1
Frank	xv_{n_v}	0	0	0	999	1
Gaussian	AIC	0	24	0	7	969
Gaussian	xv_1	0	24	0	7	969
Gaussian	xv_{CIC}	0	50	0	13	937
Gaussian	xv_{n_v}	0	30	0	7	963

Table 6: Copula selection using different information criteria ($n = 250$, $\tau = 0.75$)

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	995	0	0	0	5
Clayton	xv_1	995	0	0	0	5
Clayton	xv_{CIC}	989	0	0	2	9
Clayton	xv_{n_v}	995	0	0	0	5
Gumbel	AIC	0	818	140	9	33
Gumbel	xv_1	0	818	140	9	33
Gumbel	xv_{CIC}	0	804	160	9	27
Gumbel	xv_{n_v}	0	825	134	9	32
Joe	AIC	0	71	929	0	0
Joe	xv_1	0	71	929	0	0
Joe	xv_{CIC}	0	61	939	0	0
Joe	xv_{n_v}	0	74	926	0	0
Frank	AIC	5	10	0	868	117
Frank	xv_1	5	10	0	868	117
Frank	xv_{CIC}	4	12	0	878	106
Frank	xv_{n_v}	5	13	0	872	110
Gaussian	AIC	13	28	0	82	877
Gaussian	xv_1	13	28	0	82	877
Gaussian	xv_{CIC}	8	36	0	100	856
Gaussian	xv_{n_v}	13	31	0	92	864

Table 7: Copula selection using different information criteria ($n = 500$, $\tau = 0.25$)

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	1000	0	0	0	0
Clayton	xv_1	1000	0	0	0	0
Clayton	xv_{CIC}	1000	0	0	0	0
Clayton	xv_{n_v}	1000	0	0	0	0
Gumbel	AIC	0	981	15	1	3
Gumbel	xv_1	0	981	15	1	3
Gumbel	xv_{CIC}	0	983	15	1	1
Gumbel	xv_{n_v}	0	984	13	1	2
Joe	AIC	0	3	997	0	0
Joe	xv_1	0	3	997	0	0
Joe	xv_{CIC}	0	1	999	0	0
Joe	xv_{n_v}	0	3	997	0	0
Frank	AIC	0	0	0	995	5
Frank	xv_1	0	0	0	995	5
Frank	xv_{CIC}	0	0	0	997	3
Frank	xv_{n_v}	0	0	0	995	5
Gaussian	AIC	0	3	0	6	991
Gaussian	xv_1	0	3	0	6	991
Gaussian	xv_{CIC}	0	5	0	11	984
Gaussian	xv_{n_v}	0	3	0	6	991

Table 8: Copula selection using different information criteria ($n = 500$, $\tau = 0.50$)

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	1000	0	0	0	0
Clayton	xv_1	1000	0	0	0	0
Clayton	xv_{CIC}	1000	0	0	0	0
Clayton	xv_{n_v}	1000	0	0	0	0
Gumbel	AIC	0	998	1	0	1
Gumbel	xv_1	0	998	1	0	1
Gumbel	xv_{CIC}	0	997	2	0	1
Gumbel	xv_{n_v}	0	998	1	0	1
Joe	AIC	0	0	1000	0	0
Joe	xv_1	0	0	1000	0	0
Joe	xv_{CIC}	0	0	1000	0	0
Joe	xv_{n_v}	0	0	1000	0	0
Frank	AIC	0	0	0	1000	0
Frank	xv_1	0	0	0	1000	0
Frank	xv_{CIC}	0	0	0	1000	0
Frank	xv_{n_v}	0	0	0	1000	0
Gaussian	AIC	0	1	0	0	999
Gaussian	xv_1	0	1	0	0	999
Gaussian	xv_{CIC}	0	5	0	0	995
Gaussian	xv_{n_v}	0	2	0	0	998

Table 9: Copula selection using different information criteria ($n = 500$, $\tau = 0.75$)

4.2 $\tau = 0.05$ and $n \in \{100, 200\}$

For extremely weak dependence, $\tau = 0.05$ (when the copulas are close to the independence copula), and a small sample size of $n = 100$, Table 10 shows that in most of the 1000 replications, none of the proposed information criteria is able to correctly select the model when the true copula is Gumbel, Frank, or Gaussian. Note that our proposed xv_{n_v} fails to distinguish the Clayton copula in the majority of the 1000 replications, whereas the other information criteria are able to do so for the sample size $n = 100$.

One can see in Table 11 that increasing the sample size to $n = 200$ helps slightly for the criteria AIC and xv_1, as they correctly identify the Frank copula more often. However, the larger sample size $n = 200$ still does not help xv_{n_v}, even for selecting the Clayton copula. It is interesting to observe that when the data are generated from the Gumbel copula, xv_{n_v} still fails to select the correct model in most replications, but it chooses the Gumbel copula more often than the other methods for both $n \in \{100, 200\}$.

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	494	55	86	241	124
Clayton	xv_1	494	55	86	241	124
Clayton	$\mathrm{xv}_{\mathrm{CIC}}$	542	33	95	178	152
Clayton	xv_{n_v}	215	186	386	138	75
Gumbel	AIC	269	72	330	205	124
Gumbel	xv_1	272	71	330	206	121
Gumbel	$\mathrm{xv}_{\mathrm{CIC}}$	335	52	336	151	126
Gumbel	xv_{n_v}	87	284	441	115	73
Joe	AIC	246	84	424	156	90
Joe	xv_1	246	84	423	156	91
Joe	$\mathrm{xv}_{\mathrm{CIC}}$	316	49	434	118	83
Joe	xv_{n_v}	65	297	499	86	53
Frank	AIC	361	51	170	271	147
Frank	xv_1	361	50	171	272	146
Frank	$\mathrm{xv}_{\mathrm{CIC}}$	422	36	186	208	148
Frank	xv_{n_v}	106	242	402	164	86
Gaussian	AIC	347	62	197	250	144
Gaussian	xv_1	347	62	197	251	143
Gaussian	$\mathrm{xv}_{\mathrm{CIC}}$	402	35	207	188	168
Gaussian	xv_{n_v}	119	248	376	149	108

Table 10: Copula selection using different information criteria ($n = 100$, $\tau = 0.05$)

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	555	52	89	169	135
Clayton	xv_1	555	52	89	169	135
Clayton	xv_{CIC}	557	48	92	160	143
Clayton	xv_{n_v}	310	145	362	102	81
Gumbel	AIC	194	132	353	190	131
Gumbel	xv_1	195	132	352	190	131
Gumbel	xv_{CIC}	242	106	374	155	123
Gumbel	xv_{n_v}	92	262	436	117	93
Joe	AIC	143	118	513	132	94
Joe	xv_1	144	118	513	132	93
Joe	xv_{CIC}	187	89	536	99	89
Joe	xv_{n_v}	46	278	537	78	61
Frank	AIC	294	75	151	328	152
Frank	xv_1	294	74	152	328	152
Frank	xv_{CIC}	329	61	154	292	164
Frank	xv_{n_v}	124	213	332	230	101
Gaussian	AIC	327	86	167	240	180
Gaussian	xv_1	327	86	167	240	180
Gaussian	xv_{CIC}	370	73	174	199	184
Gaussian	xv_{n_v}	138	215	354	153	140

Table 11: Copula selection using different information criteria ($n = 200$, $\tau = 0.05$)

4.3 $\tau = 0.10$ and $n \in \{100, 200\}$

Here, we can see that for $\tau = 0.10$ and the smaller sample size $n = 100$, the information criteria still fail to correctly select the Gumbel or Gaussian copula in most of the 1000 replications. However, when the sample size is increased to $n = 200$, all criteria most often select the true Gaussian copula.

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	630	48	49	150	123
Clayton	xv_1	631	49	49	149	122
Clayton	xv_{CIC}	618	40	59	137	146
Clayton	xv_{n_v}	432	94	218	123	133
Gumbel	AIC	144	160	438	153	105
Gumbel	xv_1	144	160	438	153	105
Gumbel	xv_{CIC}	169	130	464	128	109
Gumbel	xv_{n_v}	87	259	455	113	86
Joe	AIC	87	123	627	96	67
Joe	xv_1	87	123	627	96	67
Joe	xv_{CIC}	117	96	644	79	64
Joe	xv_{n_v}	43	237	590	80	50
Frank	AIC	282	73	158	348	139
Frank	xv_1	283	73	157	348	139
Frank	xv_{CIC}	283	39	191	334	153
Frank	xv_{n_v}	120	172	282	298	128
Gaussian	AIC	302	109	188	215	186
Gaussian	xv_1	302	109	189	215	185
Gaussian	xv_{CIC}	301	84	218	186	211
Gaussian	xv_{n_v}	169	222	269	170	170

Table 12: Copula selection using different information criteria ($n = 100$, $\tau = 0.10$)

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	709	50	15	119	107
Clayton	xv_1	709	50	15	119	107
Clayton	$\mathrm{xv}_{\mathrm{CIC}}$	693	52	15	117	123
Clayton	xv_{n_v}	636	68	66	117	113
Gumbel	AIC	82	291	397	122	108
Gumbel	xv_1	82	290	397	123	108
Gumbel	$\mathrm{xv}_{\mathrm{CIC}}$	87	261	438	119	95
Gumbel	xv_{n_v}	67	347	380	113	93
Joe	AIC	22	181	696	55	46
Joe	xv_1	22	181	696	55	46
Joe	$\mathrm{xv}_{\mathrm{CIC}}$	24	144	738	50	44
Joe	xv_{n_v}	8	240	658	55	39
Frank	AIC	204	112	80	425	179
Frank	xv_1	204	112	80	425	179
Frank	$\mathrm{xv}_{\mathrm{CIC}}$	200	96	95	420	189
Frank	xv_{n_v}	154	153	135	399	159
Gaussian	AIC	239	144	109	223	285
Gaussian	xv_1	239	144	109	223	285
Gaussian	$\mathrm{xv}_{\mathrm{CIC}}$	222	131	132	215	300
Gaussian	xv_{n_v}	179	193	134	213	281

Table 13: Copula selection using different information criteria ($n = 200$, $\tau = 0.10$)

4.4 $\tau \in \{0.15, 0.20\}$ and $n \in \{100, 200\}$

For $\tau \in \{0.15, 0.20\}$, all information criteria still struggle to correctly select the Gumbel copula when the sample size is $n = 100$. Moreover, for $\tau = 0.15$ and $n = 100$, information criteria AIC and xv_1 fail to select the true Gaussian copula in the majority of replications.

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	743	45	13	105	94
Clayton	xv_1	744	45	13	104	94
Clayton	$\mathrm{xv}_{\mathrm{CIC}}$	702	42	21	111	124
Clayton	xv_{n_v}	644	66	72	111	107
Gumbel	AIC	92	263	399	119	127
Gumbel	xv_1	93	261	400	119	127
Gumbel	$\mathrm{xv}_{\mathrm{CIC}}$	94	211	453	116	126
Gumbel	xv_{n_v}	73	301	397	114	115
Joe	AIC	27	160	690	61	62
Joe	xv_1	27	160	690	61	62
Joe	$\mathrm{xv}_{\mathrm{CIC}}$	43	112	731	62	52
Joe	xv_{n_v}	18	189	666	70	57
Frank	AIC	201	105	108	400	186
Frank	xv_1	201	104	110	400	185
Frank	$\mathrm{xv}_{\mathrm{CIC}}$	178	92	140	382	208
Frank	xv_{n_v}	159	150	134	381	176
Gaussian	AIC	241	140	149	234	236
Gaussian	xv_1	241	140	149	234	236
Gaussian	$\mathrm{xv}_{\mathrm{CIC}}$	207	123	184	224	262
Gaussian	xv_{n_v}	188	185	173	218	236

Table 14: Copula selection using different information criteria ($n = 100$, $\tau = 0.15$)

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	827	17	2	76	78
Clayton	xv_1	827	17	2	76	78
Clayton	xv_{CIC}	784	17	4	85	110
Clayton	xv_{n_v}	806	20	10	85	79
Gumbel	AIC	32	379	366	87	136
Gumbel	xv_1	32	379	366	87	136
Gumbel	xv_{CIC}	28	348	415	89	120
Gumbel	xv_{n_v}	31	405	343	95	126
Joe	AIC	2	185	761	25	27
Joe	xv_1	2	185	761	25	27
Joe	xv_{CIC}	1	155	796	24	24
Joe	xv_{n_v}	1	197	753	23	26
Frank	AIC	139	118	42	509	192
Frank	xv_1	139	117	43	509	192
Frank	xv_{CIC}	109	112	58	517	204
Frank	xv_{n_v}	121	142	36	512	189
Gaussian	AIC	183	169	44	199	405
Gaussian	xv_1	183	169	44	199	405
Gaussian	xv_{CIC}	153	175	59	201	412
Gaussian	xv_{n_v}	181	178	48	203	390

Table 15: Copula selection using different information criteria ($n = 200$, $\tau = 0.15$)

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	792	26	5	68	109
Clayton	xv_1	792	26	5	67	110
Clayton	xv_{CIC}	748	33	6	73	140
Clayton	xv_{n_v}	775	37	10	75	103
Gumbel	AIC	40	336	391	95	138
Gumbel	xv_1	40	336	391	95	138
Gumbel	xv_{CIC}	31	272	463	103	131
Gumbel	xv_{n_v}	34	369	378	98	121
Joe	AIC	17	201	706	40	36
Joe	xv_1	17	201	706	40	36
Joe	xv_{CIC}	17	151	764	41	27
Joe	xv_{n_v}	12	222	697	37	32
Frank	AIC	164	119	82	458	177
Frank	xv_1	164	119	82	458	177
Frank	xv_{CIC}	120	116	115	461	188
Frank	xv_{n_v}	162	148	82	442	166
Gaussian	AIC	192	157	105	237	309
Gaussian	xv_1	192	158	104	237	309
Gaussian	xv_{CIC}	151	146	142	235	326
Gaussian	xv_{n_v}	188	191	102	239	280

Table 16: Copula selection using different information criteria ($n = 100$, $\tau = 0.20$)

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	900	7	0	36	57
Clayton	xv_1	900	7	0	36	57
Clayton	xv_{CIC}	870	7	0	48	75
Clayton	xv_{n_v}	897	7	0	39	57
Gumbel	AIC	14	496	318	74	98
Gumbel	xv_1	14	495	319	74	98
Gumbel	xv_{CIC}	11	455	372	74	88
Gumbel	xv_{n_v}	13	524	300	72	91
Joe	AIC	0	187	801	5	7
Joe	xv_1	0	187	801	5	7
Joe	xv_{CIC}	0	152	840	4	4
Joe	xv_{n_v}	0	189	798	6	7
Frank	AIC	96	106	11	591	196
Frank	xv_1	96	106	11	591	196
Frank	xv_{CIC}	70	110	16	612	192
Frank	xv_{n_v}	90	113	10	598	189
Gaussian	AIC	130	164	16	192	498
Gaussian	xv_1	130	164	16	192	498
Gaussian	xv_{CIC}	99	185	19	206	491
Gaussian	xv_{n_v}	128	184	14	201	473

Table 17: Copula selection using different information criteria ($n = 200$, $\tau = 0.20$)

5 Coincidence percentages

The following tables show the coincidence percentages (i.e., the fraction of times two methods select the same model, regardless of whether it is the true model) between the cross-validation based information criteria and AIC across all considered copula families. The estimated 95 % confidence intervals are based upon the asymptotic approximation to the standard normal distribution, which can be used due to the size of the data-sets (5000 for each non-empty cell in the τ-columns).

From the tables, one can see that xv_1 is the method most similar to AIC in the sense of coincidence percentages. The approximation method $\mathrm{xv}_{\mathrm{CIC}}$ is much closer to AIC under weak dependence, i.e., $\tau \in \{0.05, 0.10, 0.15\}$, than the proposed method xv_{n_v}.

Note that in Table 18, for $\tau \in \{0.5, 0.75\}$, most of the confidence intervals are too narrow relative to the precision used in the tables. Therefore, they are reported as ± 0.000.

	$\tau = 0.05$	$\tau = 0.1$	$\tau = 0.15$	$\tau = 0.2$	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$	All
100	99.74 ± 0.141	99.92 ± 0.078	99.90 ± 0.088	99.96 ± 0.055	99.94 ± 0.068	100.00 ± 0.000	99.98 ± 0.039	99.92 ± 0.030
200	99.92 ± 0.078	99.98 ± 0.039	99.98 ± 0.039	99.98 ± 0.039				99.97 ± 0.026
250					99.98 ± 0.039	100.00 ± 0.000	100.00 ± 0.000	99.99 ± 0.013
500					100.00 ± 0.000	100.00 ± 0.000	100.00 ± 0.000	100.00 ± 0.000

Table 18: Coincedence of AIC and xv_1 are shown with 95 % confidence intervals, and all values are expressed as percentages.

	$\tau = 0.05$	$\tau = 0.1$	$\tau = 0.15$	$\tau = 0.2$	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$	All
100	79.26 ± 1.124	86.46 ± 0.949	89.10 ± 0.864	89.34 ± 0.856	90.34 ± 0.819	93.78 ± 0.669	95.30 ± 0.587	89.08 ± 0.327
200	86.88 ± 0.936	91.38 ± 0.778	92.90 ± 0.712	94.10 ± 0.653				91.31 ± 0.390
250					95.48 ± 0.576	98.40 ± 0.348	99.08 ± 0.265	97.65 ± 0.242
500					98.16 ± 0.372	99.74 ± 0.141	99.90 ± 0.088	99.27 ± 0.136

Table 19: Coincedence of AIC and $\mathrm{xv}_{\mathrm{CIC}}$ are shown with 95 % confidence intervals, and all values are expressed as percentages.

	$\tau = 0.05$	$\tau = 0.1$	$\tau = 0.15$	$\tau = 0.2$	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$	All
100	47.76 ± 1.385	67.84 ± 1.295	80.44 ± 1.100	87.40 ± 0.920	90.74 ± 0.803	95.40 ± 0.581	97.54 ± 0.429	81.02 ± 0.411
200	59.28 ± 1.362	83.14 ± 1.038	91.56 ± 0.771	94.64 ± 0.624				82.16 ± 0.531
250					97.08 ± 0.467	98.86 ± 0.294	99.72 ± 0.146	98.55 ± 0.191
500					99.00 ± 0.276	99.86 ± 0.104	99.98 ± 0.039	99.61 ± 0.099

Table 20: Coincedence of AIC and xv_{n_v} are shown with 95 % confidence intervals, and all values are expressed as percentages.

6 Hit rates

In the following two subsections, we present tables of hit rates. By the hit rate in each cell, we mean the fraction of times a specific criterion (from the rows) selected the correct copula (from the columns), divided by the number of replications (1000). The estimated 95 % confidence intervals are based upon the asymptotic approximation to the standard normal distribution, which can be used due to the size of the data sets (1000 for each cell).

Note that regardless of the sample size and the value of Kendall's tau, the most challenging copulas to identify for all criteria are Gaussian and Gumbel. Also, in the specific case when the true copula model is Gumbel, the proposed xv_{n_v} performed better (in terms of hit rates and their confidence intervals) than other criteria for all considered values of Kendall's tau and sample sizes.

6.1 $\tau \in \{0.05, 0.10, 0.15, 0.20\}$ and $n \in \{100, 200\}$

From the tables, one can see that in the case of weak dependence, i.e., $\tau \in \{0.05, 0.10, 0.15, 0.20\}$, all information criteria perform poorly in correctly selecting the Gumbel, Frank, and Gaussian copulas.

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	49.40 ± 3.100	7.20 ± 1.600	42.40 ± 3.060	27.10 ± 2.760	14.40 ± 2.180
xv_1	49.40 ± 3.100	7.10 ± 1.590	42.30 ± 3.060	27.20 ± 2.760	14.30 ± 2.170
$\mathrm{xv}_{\mathrm{CIC}}$	54.20 ± 3.090	5.20 ± 1.380	43.40 ± 3.070	20.80 ± 2.520	16.80 ± 2.320
xv_{n_v}	21.50 ± 2.550	28.40 ± 2.800	49.90 ± 3.100	16.40 ± 2.300	10.80 ± 1.920

Table 21: Hit rates ($n = 100$, $\tau = 0.05$) are shown with 95 % confidence intervals, and all values are expressed as percentages.

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	55.50 ± 3.080	13.20 ± 2.100	51.30 ± 3.100	32.80 ± 2.910	18.00 ± 2.380
xv_1	55.50 ± 3.080	13.20 ± 2.100	51.30 ± 3.100	32.80 ± 2.910	18.00 ± 2.380
$\mathrm{xv}_{\mathrm{CIC}}$	55.70 ± 3.080	10.60 ± 1.910	53.60 ± 3.090	29.20 ± 2.820	18.40 ± 2.400
xv_{n_v}	31.00 ± 2.870	26.20 ± 2.730	53.70 ± 3.090	23.00 ± 2.610	14.00 ± 2.150

Table 22: Hit rates ($n = 200$, $\tau = 0.05$) are shown with 95 % confidence intervals, and all values are expressed as percentages.

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	63.00 ± 2.990	16.00 ± 2.270	62.70 ± 3.000	34.80 ± 2.950	18.60 ± 2.410
xv_1	63.10 ± 2.990	16.00 ± 2.270	62.70 ± 3.000	34.80 ± 2.950	18.50 ± 2.410
$\mathrm{xv}_{\mathrm{CIC}}$	61.80 ± 3.010	13.00 ± 2.090	64.40 ± 2.970	33.40 ± 2.920	21.10 ± 2.530
xv_{n_v}	43.20 ± 3.070	25.90 ± 2.720	59.00 ± 3.050	29.80 ± 2.840	17.00 ± 2.330

Table 23: Hit rates ($n = 100$, $\tau = 0.10$) are shown with 95 % confidence intervals, and all values are expressed as percentages.

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	70.90 ± 2.820	29.10 ± 2.820	69.60 ± 2.850	42.50 ± 3.070	28.50 ± 2.800
xv_1	70.90 ± 2.820	29.00 ± 2.810	69.60 ± 2.850	42.50 ± 3.070	28.50 ± 2.800
$\mathrm{xv}_{\mathrm{CIC}}$	69.30 ± 2.860	26.10 ± 2.720	73.80 ± 2.730	42.00 ± 3.060	30.00 ± 2.840
xv_{n_v}	63.60 ± 2.980	34.70 ± 2.950	65.80 ± 2.940	39.90 ± 3.040	28.10 ± 2.790

Table 24: Hit rates ($n = 200$, $\tau = 0.10$) are shown with 95 % confidence intervals, and all values are expressed as percentages.

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	74.30 ± 2.710	26.30 ± 2.730	69.00 ± 2.870	40.00 ± 3.040	23.60 ± 2.630
xv_1	74.40 ± 2.710	26.10 ± 2.720	69.00 ± 2.870	40.00 ± 3.040	23.60 ± 2.630
$\mathrm{xv}_{\mathrm{CIC}}$	70.20 ± 2.840	21.10 ± 2.530	73.10 ± 2.750	38.20 ± 3.010	26.20 ± 2.730
xv_{n_v}	64.40 ± 2.970	30.10 ± 2.840	66.60 ± 2.920	38.10 ± 3.010	23.60 ± 2.630

Table 25: Hit rates ($n = 100$, $\tau = 0.15$) are shown with 95 % confidence intervals, and all values are expressed as percentages.

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	82.70 ± 2.350	37.90 ± 3.010	76.10 ± 2.640	50.90 ± 3.100	40.50 ± 3.040
xv_1	82.70 ± 2.350	37.90 ± 3.010	76.10 ± 2.640	50.90 ± 3.100	40.50 ± 3.040
$\mathrm{xv}_{\mathrm{CIC}}$	78.40 ± 2.550	34.80 ± 2.950	79.60 ± 2.500	51.70 ± 3.100	41.20 ± 3.050
xv_{n_v}	80.60 ± 2.450	40.50 ± 3.040	75.30 ± 2.670	51.20 ± 3.100	39.00 ± 3.020

Table 26: Hit rates ($n = 200$, $\tau = 0.15$) are shown with 95 % confidence intervals, and all values are expressed as percentages.

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	79.20 ± 2.520	33.60 ± 2.930	70.60 ± 2.830	45.80 ± 3.090	30.90 ± 2.870
xv_1	79.20 ± 2.520	33.60 ± 2.930	70.60 ± 2.830	45.80 ± 3.090	30.90 ± 2.870
$\mathrm{xv}_{\mathrm{CIC}}$	74.80 ± 2.690	27.20 ± 2.760	76.40 ± 2.630	46.10 ± 3.090	32.60 ± 2.910
xv_{n_v}	77.50 ± 2.590	36.90 ± 2.990	69.70 ± 2.850	44.20 ± 3.080	28.00 ± 2.780

Table 27: Hit rates ($n = 100$, $\tau = 0.20$) are shown with 95 % confidence intervals, and all values are expressed as percentages.

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	90.00 ± 1.860	49.60 ± 3.100	80.10 ± 2.480	59.10 ± 3.050	49.80 ± 3.100
xv_1	90.00 ± 1.860	49.50 ± 3.100	80.10 ± 2.480	59.10 ± 3.050	49.80 ± 3.100
$\mathrm{xv}_{\mathrm{CIC}}$	87.00 ± 2.090	45.50 ± 3.090	84.00 ± 2.270	61.20 ± 3.020	49.10 ± 3.100
xv_{n_v}	89.70 ± 1.880	52.40 ± 3.100	79.80 ± 2.490	59.80 ± 3.040	47.30 ± 3.100

Table 28: Hit rates ($n = 200$, $\tau = 0.20$) are shown with 95 % confidence intervals, and all values are expressed as percentages.

6.2 $\tau \in \{0.25, 0.5, 0.75\}$ and $n \in \{100, 250, 500\}$

From the following tables, one can see that as dependence increases, the performance of all information criteria improves, since it becomes easier to distinguish between copulas.

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	86.00 ± 2.150	37.60 ± 3.000	77.70 ± 2.580	50.90 ± 3.100	39.70 ± 3.030
xv_1	86.00 ± 2.150	37.60 ± 3.000	77.70 ± 2.580	50.80 ± 3.100	39.80 ± 3.040
$\mathrm{xv}_{\mathrm{CIC}}$	80.70 ± 2.450	33.60 ± 2.930	82.30 ± 2.370	53.40 ± 3.090	39.00 ± 3.020
xv_{n_v}	85.10 ± 2.210	39.60 ± 3.030	75.30 ± 2.670	50.90 ± 3.100	36.00 ± 2.980

Table 29: Hit rates ($n = 100$, $\tau = 0.25$) are shown with 95 % confidence intervals, and all values are expressed as percentages.

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	97.10 ± 1.040	62.90 ± 3.000	84.70 ± 2.230	73.30 ± 2.740	65.70 ± 2.940
xv_1	97.10 ± 1.040	62.90 ± 3.000	84.70 ± 2.230	73.30 ± 2.740	65.60 ± 2.950
$\mathrm{xv}_{\mathrm{CIC}}$	95.30 ± 1.310	60.40 ± 3.030	88.20 ± 2.000	76.10 ± 2.640	64.60 ± 2.970
xv_{n_v}	97.30 ± 1.010	64.70 ± 2.960	83.50 ± 2.300	74.30 ± 2.710	63.70 ± 2.980

Table 30: Hit rates ($n = 250$, $\tau = 0.25$) are shown with 95 % confidence intervals, and all values are expressed as percentages.

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	99.50 ± 0.440	81.80 ± 2.390	92.90 ± 1.590	86.80 ± 2.100	87.70 ± 2.040
xv_1	99.50 ± 0.440	81.80 ± 2.390	92.90 ± 1.590	86.80 ± 2.100	87.70 ± 2.040
$\mathrm{xv}_{\mathrm{CIC}}$	98.90 ± 0.650	80.40 ± 2.460	93.90 ± 1.480	87.80 ± 2.030	85.60 ± 2.180
xv_{n_v}	99.50 ± 0.440	82.50 ± 2.360	92.60 ± 1.620	87.20 ± 2.070	86.40 ± 2.130

Table 31: Hit rates ($n = 500$, $\tau = 0.25$) are shown with 95 % confidence intervals, and all values are expressed as percentages.

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	96.70 ± 1.110	67.50 ± 2.900	84.00 ± 2.270	74.20 ± 2.710	71.50 ± 2.800
xv_1	96.70 ± 1.110	67.50 ± 2.900	84.00 ± 2.270	74.20 ± 2.710	71.50 ± 2.800
$\mathrm{xv}_{\mathrm{CIC}}$	92.80 ± 1.600	68.20 ± 2.890	87.70 ± 2.040	78.30 ± 2.560	65.10 ± 2.960
xv_{n_v}	97.00 ± 1.060	68.80 ± 2.870	83.20 ± 2.320	74.20 ± 2.710	67.20 ± 2.910

Table 32: Hit rates ($n = 100$, $\tau = 0.50$) are shown with 95 % confidence intervals, and all values are expressed as percentages.

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	99.80 ± 0.280	89.50 ± 1.900	96.50 ± 1.140	94.00 ± 1.470	90.30 ± 1.840
xv_1	99.80 ± 0.280	89.50 ± 1.900	96.50 ± 1.140	94.00 ± 1.470	90.30 ± 1.840
$\mathrm{xv}_{\mathrm{CIC}}$	99.70 ± 0.340	89.70 ± 1.880	97.30 ± 1.010	95.20 ± 1.330	87.90 ± 2.020
xv_{n_v}	99.80 ± 0.280	90.20 ± 1.840	96.30 ± 1.170	95.20 ± 1.330	89.60 ± 1.890

Table 33: Hit rates ($n = 250$, $\tau = 0.50$) are shown with 95 % confidence intervals, and all values are expressed as percentages.

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	100.00 ± 0.000	98.10 ± 0.850	99.70 ± 0.340	99.50 ± 0.440	99.10 ± 0.590
xv_1	100.00 ± 0.000	98.10 ± 0.850	99.70 ± 0.340	99.50 ± 0.440	99.10 ± 0.590
$\mathrm{xv}_{\mathrm{CIC}}$	100.00 ± 0.000	98.30 ± 0.800	99.90 ± 0.200	99.70 ± 0.340	98.40 ± 0.780
xv_{n_v}	100.00 ± 0.000	98.40 ± 0.780	99.70 ± 0.340	99.50 ± 0.440	99.10 ± 0.590

Table 34: Hit rates ($n = 500$, $\tau = 0.50$) are shown with 95 % confidence intervals, and all values are expressed as percentages.

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	99.00 ± 0.620	80.60 ± 2.450	91.20 ± 1.760	89.50 ± 1.900	82.10 ± 2.380
xv_1	99.00 ± 0.620	80.60 ± 2.450	91.20 ± 1.760	89.40 ± 1.910	82.10 ± 2.380
$\mathrm{xv}_{\mathrm{CIC}}$	97.30 ± 1.010	82.40 ± 2.360	93.40 ± 1.540	92.70 ± 1.610	74.70 ± 2.700
xv_{n_v}	98.90 ± 0.650	82.20 ± 2.370	90.20 ± 1.840	89.90 ± 1.870	80.00 ± 2.480

Table 35: Hit rates ($n = 100$, $\tau = 0.75$) are shown with 95 % confidence intervals, and all values are expressed as percentages.

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	100.00 ± 0.000	96.90 ± 1.070	99.60 ± 0.390	99.90 ± 0.200	96.90 ± 1.070
xv_1	100.00 ± 0.000	96.90 ± 1.070	99.60 ± 0.390	99.90 ± 0.200	96.90 ± 1.070
$\mathrm{xv}_{\mathrm{CIC}}$	100.00 ± 0.000	97.80 ± 0.910	99.60 ± 0.390	99.90 ± 0.200	93.70 ± 1.510
xv_{n_v}	100.00 ± 0.000	97.00 ± 1.060	99.60 ± 0.390	99.90 ± 0.200	96.30 ± 1.170

Table 36: Hit rates ($n = 250$, $\tau = 0.75$) are shown with 95 % confidence intervals, and all values are expressed as percentages.

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	100.00 ± 0.000	99.80 ± 0.280	100.00 ± 0.000	100.00 ± 0.000	99.90 ± 0.200
xv_1	100.00 ± 0.000	99.80 ± 0.280	100.00 ± 0.000	100.00 ± 0.000	99.90 ± 0.200
$\mathrm{xv}_{\mathrm{CIC}}$	100.00 ± 0.000	99.70 ± 0.340	100.00 ± 0.000	100.00 ± 0.000	99.50 ± 0.440
xv_{n_v}	100.00 ± 0.000	99.80 ± 0.280	100.00 ± 0.000	100.00 ± 0.000	99.80 ± 0.280

Table 37: Hit rates ($n = 500$, $\tau = 0.75$) are shown with 95 % confidence intervals, and all values are expressed as percentages.