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 $\{1, \dots, n\}$, 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 5000 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 5000 replications.

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

Here, in most of the 5000 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	4256	73	3	273	395
Clayton	xv_1	4256	73	3	273	395
Clayton	xv_CIC	3966	79	3	360	592
Clayton	xv_{n_v}	4195	88	14	309	394
Gumbel	AIC	135	1979	1806	450	630
Gumbel	xv_1	135	1979	1806	450	630
Gumbel	xv_CIC	92	1756	2142	468	542
Gumbel	xv_{n_v}	146	2077	1739	496	542
Joe	AIC	6	903	3859	113	119
Joe	xv_1	6	903	3859	113	119
Joe	xv_CIC	4	703	4085	121	87
Joe	xv_{n_v}	8	990	3775	116	111
Frank	AIC	640	623	182	2559	996
Frank	xv_1	641	621	182	2558	998
Frank	xv_CIC	427	594	307	2711	961
Frank	xv_{n_v}	632	709	178	2572	909
Gaussian	AIC	846	914	285	970	1985
Gaussian	xv_1	847	914	285	969	1985
Gaussian	xv_CIC	588	950	406	1087	1969
Gaussian	xv_{n_v}	848	1015	268	1038	1831

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

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	4828	2	0	64	106
Clayton	xv_1	4828	2	0	64	106
Clayton	xv_CIC	4740	3	0	81	176
Clayton	xv_{n_v}	4835	2	0	66	97
Gumbel	AIC	5	3265	1180	189	361
Gumbel	xv_1	5	3265	1180	189	361
Gumbel	xv_CIC	4	3105	1401	198	292
Gumbel	xv_{n_v}	6	3318	1144	194	338
Joe	AIC	0	698	4284	12	6
Joe	xv_1	0	698	4284	12	6
Joe	xv_CIC	0	560	4425	12	3
Joe	xv_{n_v}	0	743	4241	10	6
Frank	AIC	163	281	7	3629	920
Frank	xv_1	163	281	7	3629	920
Frank	xv_CIC	87	308	9	3753	843
Frank	xv_{n_v}	163	305	6	3687	839
Gaussian	AIC	328	623	8	746	3295
Gaussian	xv_1	329	623	8	746	3294
Gaussian	xv_CIC	217	733	19	840	3191
Gaussian	xv_{n_v}	331	673	6	813	3177

Table 2: Copula selection using different information criteria ($n = 250$, $\tau = 0.25$)

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	4971	0	0	7	22
Clayton	xv_1	4971	0	0	7	22
Clayton	xv_CIC	4950	0	0	10	40
Clayton	xv_{n_v}	4971	0	0	6	23
Gumbel	AIC	0	4071	733	45	151
Gumbel	xv_1	0	4071	733	45	151
Gumbel	xv_CIC	0	3993	836	45	126
Gumbel	xv_{n_v}	0	4102	706	46	146
Joe	AIC	0	342	4658	0	0
Joe	xv_1	0	342	4658	0	0
Joe	xv_CIC	0	290	4710	0	0
Joe	xv_{n_v}	0	361	4639	0	0
Frank	AIC	21	62	0	4336	581
Frank	xv_1	21	62	0	4336	581
Frank	xv_CIC	10	72	0	4394	524
Frank	xv_{n_v}	22	67	0	4365	546
Gaussian	AIC	60	211	0	452	4277
Gaussian	xv_1	60	211	0	452	4277
Gaussian	xv_CIC	36	250	0	510	4204
Gaussian	xv_{n_v}	60	234	0	482	4224

Table 3: Copula selection using different information criteria ($n = 500$, $\tau = 0.25$)

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	4842	0	0	63	95
Clayton	xv_1	4842	0	0	63	95
Clayton	xv_CIC	4652	0	0	140	208
Clayton	xv_{n_v}	4828	0	0	82	90
Gumbel	AIC	9	3390	854	170	577
Gumbel	xv_1	9	3390	854	170	577
Gumbel	xv_CIC	2	3398	1029	210	361
Gumbel	xv_{n_v}	10	3464	842	186	498
Joe	AIC	0	733	4241	16	10
Joe	xv_1	0	733	4241	16	10
Joe	xv_CIC	0	570	4403	21	6
Joe	xv_{n_v}	0	755	4214	22	9
Frank	AIC	101	328	12	3664	895
Frank	xv_1	101	328	12	3664	895
Frank	xv_CIC	48	384	18	3880	670
Frank	xv_{n_v}	98	376	8	3713	805
Gaussian	AIC	228	829	15	438	3490
Gaussian	xv_1	228	829	15	437	3491
Gaussian	xv_CIC	102	1100	26	559	3213
Gaussian	xv_{n_v}	241	934	13	510	3302

Table 4: Copula selection using different information criteria ($n = 100$, $\tau = 0.50$)

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	4993	0	0	5	2
Clayton	xv_1	4993	0	0	5	2
Clayton	xv_CIC	4984	0	0	10	6
Clayton	xv_{n_v}	4993	0	0	6	1
Gumbel	AIC	0	4511	278	31	180
Gumbel	xv_1	0	4511	278	31	180
Gumbel	xv_CIC	0	4515	336	34	115
Gumbel	xv_{n_v}	0	4541	268	36	155
Joe	AIC	0	201	4799	0	0
Joe	xv_1	0	201	4799	0	0
Joe	xv_CIC	0	160	4840	0	0
Joe	xv_{n_v}	0	223	4777	0	0
Frank	AIC	3	40	0	4691	266
Frank	xv_1	3	40	0	4691	266
Frank	xv_CIC	0	40	0	4773	187
Frank	xv_{n_v}	3	41	0	4732	224
Gaussian	AIC	16	251	0	147	4586
Gaussian	xv_1	16	251	0	147	4586
Gaussian	xv_CIC	6	367	0	199	4428
Gaussian	xv_{n_v}	15	279	0	159	4547

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

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	5000	0	0	0	0
Clayton	xv_1	5000	0	0	0	0
Clayton	xv_CIC	4999	0	0	0	1
Clayton	xv_{n_v}	5000	0	0	0	0
Gumbel	AIC	0	4899	56	1	44
Gumbel	xv_1	0	4899	56	1	44
Gumbel	xv_CIC	0	4905	66	3	26
Gumbel	xv_{n_v}	0	4913	50	1	36
Joe	AIC	0	21	4979	0	0
Joe	xv_1	0	21	4979	0	0
Joe	xv_CIC	0	14	4986	0	0
Joe	xv_{n_v}	0	23	4977	0	0
Frank	AIC	0	1	0	4966	33
Frank	xv_1	0	1	0	4966	33
Frank	xv_CIC	0	1	0	4977	22
Frank	xv_{n_v}	0	1	0	4972	27
Gaussian	AIC	0	35	0	25	4940
Gaussian	xv_1	0	35	0	25	4940
Gaussian	xv_CIC	0	53	0	37	4910
Gaussian	xv_{n_v}	0	35	0	26	4939

Table 6: Copula selection using different information criteria ($n = 500$, $\tau = 0.50$)

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	4937	0	0	46	17
Clayton	xv_1	4937	0	0	46	17
Clayton	xv_CIC	4838	0	0	120	42
Clayton	xv_{n_v}	4927	0	0	56	17
Gumbel	AIC	0	3982	310	117	591
Gumbel	xv_1	0	3982	310	117	591
Gumbel	xv_CIC	0	4128	384	150	338
Gumbel	xv_{n_v}	0	4059	304	120	517
Joe	AIC	0	444	4545	11	0
Joe	xv_1	0	444	4545	11	0
Joe	xv_CIC	0	301	4684	15	0
Joe	xv_{n_v}	0	497	4490	13	0
Frank	AIC	15	114	0	4526	345
Frank	xv_1	15	114	0	4524	347
Frank	xv_CIC	8	126	0	4659	207
Frank	xv_{n_v}	15	133	0	4542	310
Gaussian	AIC	45	634	0	213	4108
Gaussian	xv_1	45	632	0	213	4110
Gaussian	xv_CIC	28	966	0	334	3672
Gaussian	xv_{n_v}	51	724	0	255	3970

Table 7: Copula selection using different information criteria ($n = 100$, $\tau = 0.75$)

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	4999	0	0	1	0
Clayton	xv_1	4999	0	0	1	0
Clayton	xv_CIC	4997	0	0	3	0
Clayton	xv_{n_v}	4999	0	0	1	0
Gumbel	AIC	0	4857	27	4	112
Gumbel	xv_1	0	4857	27	4	112
Gumbel	xv_CIC	0	4889	39	7	65
Gumbel	xv_{n_v}	0	4868	25	6	101
Joe	AIC	0	39	4961	0	0
Joe	xv_1	0	39	4961	0	0
Joe	xv_CIC	0	21	4979	0	0
Joe	xv_{n_v}	0	42	4958	0	0
Frank	AIC	0	1	0	4985	14
Frank	xv_1	0	1	0	4985	14
Frank	xv_CIC	0	2	0	4988	10
Frank	xv_{n_v}	0	1	0	4985	14
Gaussian	AIC	0	137	0	20	4843
Gaussian	xv_1	0	137	0	20	4843
Gaussian	xv_CIC	0	253	0	44	4703
Gaussian	xv_{n_v}	0	158	0	21	4821

Table 8: Copula selection using different information criteria ($n = 250$, $\tau = 0.75$)

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	5000	0	0	0	0
Clayton	xv_1	5000	0	0	0	0
Clayton	xv_CIC	5000	0	0	0	0
Clayton	xv_{n_v}	5000	0	0	0	0
Gumbel	AIC	0	4993	1	0	6
Gumbel	xv_1	0	4993	1	0	6
Gumbel	xv_CIC	0	4993	3	0	4
Gumbel	xv_{n_v}	0	4993	1	0	6
Joe	AIC	0	0	5000	0	0
Joe	xv_1	0	0	5000	0	0
Joe	xv_CIC	0	0	5000	0	0
Joe	xv_{n_v}	0	0	5000	0	0
Frank	AIC	0	0	0	5000	0
Frank	xv_1	0	0	0	5000	0
Frank	xv_CIC	0	0	0	5000	0
Frank	xv_{n_v}	0	0	0	5000	0
Gaussian	AIC	0	4	0	1	4995
Gaussian	xv_1	0	4	0	1	4995
Gaussian	xv_CIC	0	23	0	1	4976
Gaussian	xv_{n_v}	0	5	0	1	4994

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 5000 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 5000 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 $n = 200$ doesn't help. 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	2446	229	553	1116	656
Clayton	xv_1	2450	230	551	1114	655
Clayton	xv_CIC	2705	150	619	833	693
Clayton	xv_{n_v}	1049	994	1983	588	386
Gumbel	AIC	1387	403	1612	994	604
Gumbel	xv_1	1392	402	1612	993	601
Gumbel	xv_CIC	1732	268	1680	714	606
Gumbel	xv_{n_v}	427	1455	2196	561	361
Joe	AIC	1171	428	2128	768	505
Joe	xv_1	1172	427	2128	768	505
Joe	xv_CIC	1497	288	2147	585	483
Joe	xv_{n_v}	329	1400	2498	468	305
Frank	AIC	1744	273	907	1324	752
Frank	xv_1	1743	274	906	1325	752
Frank	xv_CIC	2083	185	956	1030	746
Frank	xv_{n_v}	578	1195	2011	787	429
Gaussian	AIC	1820	316	904	1173	787
Gaussian	xv_1	1825	315	904	1171	785
Gaussian	xv_CIC	2071	187	1004	900	838
Gaussian	xv_{n_v}	632	1234	1960	660	514

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

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	2715	258	396	983	648
Clayton	xv_1	2716	258	396	982	648
Clayton	xv_CIC	2805	219	440	855	681
Clayton	xv_{n_v}	1526	680	1701	644	449
Gumbel	AIC	991	673	1762	946	628
Gumbel	xv_1	993	673	1761	945	628
Gumbel	xv_CIC	1202	534	1839	816	609
Gumbel	xv_{n_v}	426	1416	2105	603	450
Joe	AIC	659	639	2632	618	452
Joe	xv_1	658	638	2632	620	452
Joe	xv_CIC	847	502	2697	514	440
Joe	xv_{n_v}	224	1339	2717	393	327
Frank	AIC	1546	356	747	1541	810
Frank	xv_1	1546	357	747	1541	809
Frank	xv_CIC	1753	274	793	1355	825
Frank	xv_{n_v}	650	1015	1773	1040	522
Gaussian	AIC	1675	387	845	1229	864
Gaussian	xv_1	1676	387	845	1228	864
Gaussian	xv_CIC	1835	334	880	1021	930
Gaussian	xv_{n_v}	719	1134	1752	746	649

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 5000 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	3126	250	253	755	616
Clayton	xv_1	3129	251	253	754	613
Clayton	xv_CIC	3037	213	314	708	728
Clayton	xv_{n_v}	2193	523	1048	662	574
Gumbel	AIC	820	808	2052	781	539
Gumbel	xv_1	821	808	2052	780	539
Gumbel	xv_CIC	928	620	2231	671	550
Gumbel	xv_{n_v}	459	1345	2166	618	412
Joe	AIC	425	640	3095	479	361
Joe	xv_1	426	638	3097	478	361
Joe	xv_CIC	536	466	3244	405	349
Joe	xv_{n_v}	188	1143	2980	408	281
Frank	AIC	1428	391	752	1682	747
Frank	xv_1	1434	390	750	1681	745
Frank	xv_CIC	1437	268	897	1540	858
Frank	xv_{n_v}	762	893	1299	1405	641
Gaussian	AIC	1526	531	922	1113	908
Gaussian	xv_1	1529	531	923	1110	907
Gaussian	xv_CIC	1512	388	1089	969	1042
Gaussian	xv_{n_v}	887	1059	1361	854	839

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

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	3571	218	79	610	522
Clayton	xv_1	3571	218	79	610	522
Clayton	xv_CIC	3467	218	90	625	600
Clayton	xv_{n_v}	3178	293	388	601	540
Gumbel	AIC	407	1325	2005	647	616
Gumbel	xv_1	407	1323	2006	648	616
Gumbel	xv_CIC	419	1183	2190	629	579
Gumbel	xv_{n_v}	302	1611	1921	604	562
Joe	AIC	111	990	3395	293	211
Joe	xv_1	111	990	3395	293	211
Joe	xv_CIC	141	806	3577	282	194
Joe	xv_{n_v}	71	1202	3246	292	189
Frank	AIC	1052	506	431	2115	896
Frank	xv_1	1051	506	431	2116	896
Frank	xv_CIC	1006	464	498	2078	954
Frank	xv_{n_v}	761	774	677	1956	832
Gaussian	AIC	1227	678	537	1157	1401
Gaussian	xv_1	1227	678	537	1157	1401
Gaussian	xv_CIC	1154	633	628	1125	1460
Gaussian	xv_{n_v}	945	915	719	1065	1356

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 $n = 100$. For $n = 200$ and $\tau = 0.15$, only $\mathrm{xv}_{\mathrm{CIC}}$ fails to select the Gumbel copula in the majority of replications. 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	3690	242	69	513	486
Clayton	xv_1	3691	242	68	513	486
Clayton	xv_CIC	3452	223	101	563	661
Clayton	xv_{n_v}	3209	324	368	585	514
Gumbel	AIC	460	1229	2075	643	593
Gumbel	xv_1	460	1225	2079	643	593
Gumbel	xv_CIC	443	980	2372	638	567
Gumbel	xv_{n_v}	384	1424	2046	613	533
Joe	AIC	132	798	3500	302	268
Joe	xv_1	132	798	3500	302	268
Joe	xv_CIC	161	592	3729	284	234
Joe	xv_{n_v}	89	979	3373	308	251
Frank	AIC	1087	524	533	1950	906
Frank	xv_1	1088	524	535	1949	904
Frank	xv_CIC	942	434	699	1924	1001
Frank	xv_{n_v}	859	743	689	1851	858
Gaussian	AIC	1300	690	717	1063	1230
Gaussian	xv_1	1303	688	718	1063	1228
Gaussian	xv_CIC	1119	625	863	1071	1322
Gaussian	xv_{n_v}	1046	935	824	1040	1155

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

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	4127	82	5	349	437
Clayton	xv_1	4127	82	5	349	437
Clayton	xv_CIC	3947	86	9	392	566
Clayton	xv_{n_v}	4058	94	35	382	431
Gumbel	AIC	169	1931	1854	453	593
Gumbel	xv_1	169	1930	1855	453	593
Gumbel	xv_CIC	137	1766	2089	455	553
Gumbel	xv_{n_v}	159	2047	1759	471	564
Joe	AIC	8	992	3755	126	119
Joe	xv_1	8	991	3756	126	119
Joe	xv_CIC	9	825	3945	119	102
Joe	xv_{n_v}	5	1084	3682	121	108
Frank	AIC	679	565	203	2584	969
Frank	xv_1	679	564	204	2584	969
Frank	xv_CIC	533	528	292	2651	996
Frank	xv_{n_v}	631	627	204	2592	946
Gaussian	AIC	945	830	236	1083	1906
Gaussian	xv_1	945	830	236	1083	1906
Gaussian	xv_CIC	784	826	306	1134	1950
Gaussian	xv_{n_v}	904	888	239	1109	1860

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

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	4024	115	21	373	467
Clayton	xv_1	4024	115	21	371	469
Clayton	xv_CIC	3734	136	26	443	661
Clayton	xv_{n_v}	3879	138	86	418	479
Gumbel	AIC	238	1587	1988	548	639
Gumbel	xv_1	239	1587	1988	549	637
Gumbel	xv_CIC	189	1355	2304	565	587
Gumbel	xv_{n_v}	243	1733	1906	566	552
Joe	AIC	52	884	3685	182	197
Joe	xv_1	52	885	3685	182	196
Joe	xv_CIC	51	667	3942	181	159
Joe	xv_{n_v}	38	1032	3568	187	175
Frank	AIC	832	611	370	2229	958
Frank	xv_1	832	612	369	2230	957
Frank	xv_CIC	612	540	547	2320	981
Frank	xv_{n_v}	783	693	397	2242	885
Gaussian	AIC	1051	813	449	1072	1615
Gaussian	xv_1	1051	814	448	1072	1615
Gaussian	xv_CIC	830	806	598	1110	1656
Gaussian	xv_{n_v}	1027	933	463	1092	1485

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

d.cop	IC	Clayton	Gumbel	Joe	Frank	Gaussian
Clayton	AIC	4532	28	0	182	258
Clayton	xv_1	4532	28	0	182	258
Clayton	xv_CIC	4361	32	0	232	375
Clayton	xv_{n_v}	4510	28	0	209	253
Gumbel	AIC	60	2477	1590	337	536
Gumbel	xv_1	60	2476	1591	337	536
Gumbel	xv_CIC	48	2319	1825	342	466
Gumbel	xv_{n_v}	64	2575	1511	351	499
Joe	AIC	2	930	3969	46	53
Joe	xv_1	2	930	3969	46	53
Joe	xv_CIC	1	748	4170	45	36
Joe	xv_{n_v}	2	1019	3886	47	46
Frank	AIC	429	500	63	3031	977
Frank	xv_1	429	500	63	3031	977
Frank	xv_CIC	267	506	104	3162	961
Frank	xv_{n_v}	407	528	56	3093	916
Gaussian	AIC	667	793	112	966	2462
Gaussian	xv_1	667	793	112	966	2462
Gaussian	xv_CIC	493	859	149	1038	2461
Gaussian	xv_{n_v}	670	866	100	1013	2351

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}. However, for greater sample sizes, $n \in \{250, 500\}$, and for greater values of Kendall's tau, $\tau \in \{0.25, 0.5, 0.75\}$, the proposed method xv_{n_v} is closer to AIC than $\mathrm{xv}_{\mathrm{CIC}}$.

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	99.77 ± 0.060	99.90 ± 0.038	99.92 ± 0.036	99.94 ± 0.029	99.97 ± 0.021	100.00 ± 0.008	99.98 ± 0.016	99.93 ± 0.013
200	99.93 ± 0.033	99.99 ± 0.014	99.99 ± 0.014	100.00 ± 0.008				99.97 ± 0.010
250					100.00 ± 0.008	100.00 ± 0.000	100.00 ± 0.000	100.00 ± 0.003
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, with 95 % confidence intervals (all values multiplied by 100).

	$\tau = 0.05$	$\tau = 0.1$	$\tau = 0.15$	$\tau = 0.2$	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$	All
100	79.85 ± 0.497	85.63 ± 0.435	88.52 ± 0.395	89.74 ± 0.376	90.25 ± 0.368	93.23 ± 0.312	95.00 ± 0.270	88.89 ± 0.147
200	86.10 ± 0.429	91.58 ± 0.344	93.09 ± 0.314	93.75 ± 0.300				91.13 ± 0.176
250					95.50 ± 0.257	98.20 ± 0.165	99.11 ± 0.117	97.60 ± 0.110
500					98.28 ± 0.161	99.68 ± 0.070	99.91 ± 0.038	99.29 ± 0.060

Table 19: Coincedence of AIC and $\mathrm{xv}_{\mathrm{CIC}}$, with 95 % confidence intervals (all values multiplied by 100).

	$\tau = 0.05$	$\tau = 0.1$	$\tau = 0.15$	$\tau = 0.2$	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$	All
100	47.86 ± 0.619	67.71 ± 0.580	80.67 ± 0.490	87.44 ± 0.411	90.44 ± 0.365	95.31 ± 0.262	97.46 ± 0.195	80.98 ± 0.184
200	59.30 ± 0.609	83.14 ± 0.464	91.77 ± 0.341	94.65 ± 0.279				82.21 ± 0.237
250					96.87 ± 0.216	99.09 ± 0.118	99.79 ± 0.057	98.58 ± 0.085
500					98.95 ± 0.126	99.88 ± 0.042	100.00 ± 0.008	99.61 ± 0.045

Table 20: Coincedence of AIC and xv_{n_v}, with 95 % confidence intervals (all values multiplied by 100).

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 (5000). 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 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	48.92 ± 1.390	8.06 ± 0.750	42.56 ± 1.370	26.48 ± 1.220	15.74 ± 1.010
xv_1	49.00 ± 1.390	8.04 ± 0.750	42.56 ± 1.370	26.50 ± 1.220	15.70 ± 1.010
$\mathrm{xv}_{\mathrm{CIC}}$	54.10 ± 1.380	5.36 ± 0.620	42.94 ± 1.370	20.60 ± 1.120	16.76 ± 1.040
xv_{n_v}	20.98 ± 1.130	29.10 ± 1.260	49.96 ± 1.390	15.74 ± 1.010	10.28 ± 0.840

Table 21: Hit rates ($n = 100$, $\tau = 0.05$), with 95 % confidence intervals (all values multiplied by 100).

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	54.30 ± 1.380	13.46 ± 0.950	52.64 ± 1.380	30.82 ± 1.280	17.28 ± 1.050
xv_1	54.32 ± 1.380	13.46 ± 0.950	52.64 ± 1.380	30.82 ± 1.280	17.28 ± 1.050
$\mathrm{xv}_{\mathrm{CIC}}$	56.10 ± 1.380	10.68 ± 0.860	53.94 ± 1.380	27.10 ± 1.230	18.60 ± 1.080
xv_{n_v}	30.52 ± 1.280	28.32 ± 1.250	54.34 ± 1.380	20.80 ± 1.130	12.98 ± 0.930

Table 22: Hit rates ($n = 200$, $\tau = 0.05$), with 95 % confidence intervals (all values multiplied by 100).

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	62.52 ± 1.340	16.16 ± 1.020	61.90 ± 1.350	33.64 ± 1.310	18.16 ± 1.070
xv_1	62.58 ± 1.340	16.16 ± 1.020	61.94 ± 1.350	33.62 ± 1.310	18.14 ± 1.070
$\mathrm{xv}_{\mathrm{CIC}}$	60.74 ± 1.350	12.40 ± 0.910	64.88 ± 1.320	30.80 ± 1.280	20.84 ± 1.130
xv_{n_v}	43.86 ± 1.380	26.90 ± 1.230	59.60 ± 1.360	28.10 ± 1.250	16.78 ± 1.040

Table 23: Hit rates ($n = 100$, $\tau = 0.10$), with 95 % confidence intervals (all values multiplied by 100).

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	71.42 ± 1.250	26.50 ± 1.220	67.90 ± 1.290	42.30 ± 1.370	28.02 ± 1.240
xv_1	71.42 ± 1.250	26.46 ± 1.220	67.90 ± 1.290	42.32 ± 1.370	28.02 ± 1.240
$\mathrm{xv}_{\mathrm{CIC}}$	69.34 ± 1.280	23.66 ± 1.180	71.54 ± 1.250	41.56 ± 1.370	29.20 ± 1.260
xv_{n_v}	63.56 ± 1.330	32.22 ± 1.300	64.92 ± 1.320	39.12 ± 1.350	27.12 ± 1.230

Table 24: Hit rates ($n = 200$, $\tau = 0.10$), with 95 % confidence intervals (all values multiplied by 100).

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	73.80 ± 1.220	24.58 ± 1.190	70.00 ± 1.270	39.00 ± 1.350	24.60 ± 1.190
xv_1	73.82 ± 1.220	24.50 ± 1.190	70.00 ± 1.270	38.98 ± 1.350	24.56 ± 1.190
$\mathrm{xv}_{\mathrm{CIC}}$	69.04 ± 1.280	19.60 ± 1.100	74.58 ± 1.210	38.48 ± 1.350	26.44 ± 1.220
xv_{n_v}	64.18 ± 1.330	28.48 ± 1.250	67.46 ± 1.300	37.02 ± 1.340	23.10 ± 1.170

Table 25: Hit rates ($n = 100$, $\tau = 0.15$), with 95 % confidence intervals (all values multiplied by 100).

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	82.54 ± 1.050	38.62 ± 1.350	75.10 ± 1.200	51.68 ± 1.390	38.12 ± 1.350
xv_1	82.54 ± 1.050	38.60 ± 1.350	75.12 ± 1.200	51.68 ± 1.390	38.12 ± 1.350
$\mathrm{xv}_{\mathrm{CIC}}$	78.94 ± 1.130	35.32 ± 1.320	78.90 ± 1.130	53.02 ± 1.380	39.00 ± 1.350
xv_{n_v}	81.16 ± 1.080	40.94 ± 1.360	73.64 ± 1.220	51.84 ± 1.390	37.20 ± 1.340

Table 26: Hit rates ($n = 200$, $\tau = 0.15$), with 95 % confidence intervals (all values multiplied by 100).

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	80.48 ± 1.100	31.74 ± 1.290	73.70 ± 1.220	44.58 ± 1.380	32.30 ± 1.300
xv_1	80.48 ± 1.100	31.74 ± 1.290	73.70 ± 1.220	44.60 ± 1.380	32.30 ± 1.300
$\mathrm{xv}_{\mathrm{CIC}}$	74.68 ± 1.210	27.10 ± 1.230	78.84 ± 1.130	46.40 ± 1.380	33.12 ± 1.300
xv_{n_v}	77.58 ± 1.160	34.66 ± 1.320	71.36 ± 1.250	44.84 ± 1.380	29.70 ± 1.270

Table 27: Hit rates ($n = 100$, $\tau = 0.20$), with 95 % confidence intervals (all values multiplied by 100).

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	90.64 ± 0.810	49.54 ± 1.390	79.38 ± 1.120	60.62 ± 1.350	49.24 ± 1.390
xv_1	90.64 ± 0.810	49.52 ± 1.390	79.38 ± 1.120	60.62 ± 1.350	49.24 ± 1.390
$\mathrm{xv}_{\mathrm{CIC}}$	87.22 ± 0.930	46.38 ± 1.380	83.40 ± 1.030	63.24 ± 1.340	49.22 ± 1.390
xv_{n_v}	90.20 ± 0.820	51.50 ± 1.390	77.72 ± 1.150	61.86 ± 1.350	47.02 ± 1.380

Table 28: Hit rates ($n = 200$, $\tau = 0.20$), with 95 % confidence intervals (all values multiplied by 100).

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	85.12 ± 0.990	39.58 ± 1.360	77.18 ± 1.160	51.18 ± 1.390	39.70 ± 1.360
xv_1	85.12 ± 0.990	39.58 ± 1.360	77.18 ± 1.160	51.16 ± 1.390	39.70 ± 1.360
$\mathrm{xv}_{\mathrm{CIC}}$	79.32 ± 1.120	35.12 ± 1.320	81.70 ± 1.070	54.22 ± 1.380	39.38 ± 1.350
xv_{n_v}	83.90 ± 1.020	41.54 ± 1.370	75.50 ± 1.190	51.44 ± 1.390	36.62 ± 1.340

Table 29: Hit rates ($n = 100$, $\tau = 0.25$), with 95 % confidence intervals (all values multiplied by 100).

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	96.56 ± 0.510	65.30 ± 1.320	85.68 ± 0.970	72.58 ± 1.240	65.90 ± 1.310
xv_1	96.56 ± 0.510	65.30 ± 1.320	85.68 ± 0.970	72.58 ± 1.240	65.88 ± 1.310
$\mathrm{xv}_{\mathrm{CIC}}$	94.80 ± 0.620	62.10 ± 1.340	88.50 ± 0.880	75.06 ± 1.200	63.82 ± 1.330
xv_{n_v}	96.70 ± 0.500	66.36 ± 1.310	84.82 ± 0.990	73.74 ± 1.220	63.54 ± 1.330

Table 30: Hit rates ($n = 250$, $\tau = 0.25$), with 95 % confidence intervals (all values multiplied by 100).

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	99.42 ± 0.210	81.42 ± 1.080	93.16 ± 0.700	86.72 ± 0.940	85.54 ± 0.970
xv_1	99.42 ± 0.210	81.42 ± 1.080	93.16 ± 0.700	86.72 ± 0.940	85.54 ± 0.970
$\mathrm{xv}_{\mathrm{CIC}}$	99.00 ± 0.280	79.86 ± 1.110	94.20 ± 0.650	87.88 ± 0.900	84.08 ± 1.010
xv_{n_v}	99.42 ± 0.210	82.04 ± 1.060	92.78 ± 0.720	87.30 ± 0.920	84.48 ± 1.000

Table 31: Hit rates ($n = 500$, $\tau = 0.25$), with 95 % confidence intervals (all values multiplied by 100).

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	96.84 ± 0.480	67.80 ± 1.300	84.82 ± 0.990	73.28 ± 1.230	69.80 ± 1.270
xv_1	96.84 ± 0.480	67.80 ± 1.300	84.82 ± 0.990	73.28 ± 1.230	69.82 ± 1.270
xv_{CIC}	93.04 ± 0.710	67.96 ± 1.290	88.06 ± 0.900	77.60 ± 1.160	64.26 ± 1.330
xv_{n_v}	96.56 ± 0.510	69.28 ± 1.280	84.28 ± 1.010	74.26 ± 1.210	66.04 ± 1.310

Table 32: Hit rates ($n = 100$, $\tau = 0.50$), with 95 % confidence intervals (all values multiplied by 100).

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	99.86 ± 0.100	90.22 ± 0.820	95.98 ± 0.540	93.82 ± 0.670	91.72 ± 0.760
xv_1	99.86 ± 0.100	90.22 ± 0.820	95.98 ± 0.540	93.82 ± 0.670	91.72 ± 0.760
xv_{CIC}	99.68 ± 0.160	90.30 ± 0.820	96.80 ± 0.490	95.46 ± 0.580	88.56 ± 0.880
xv_{n_v}	99.86 ± 0.100	90.82 ± 0.800	95.54 ± 0.570	94.64 ± 0.620	90.94 ± 0.800

Table 33: Hit rates ($n = 250$, $\tau = 0.50$), with 95 % confidence intervals (all values multiplied by 100).

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	100.00 ± 0.000	97.98 ± 0.390	99.58 ± 0.180	99.32 ± 0.230	98.80 ± 0.300
xv_1	100.00 ± 0.000	97.98 ± 0.390	99.58 ± 0.180	99.32 ± 0.230	98.80 ± 0.300
xv_{CIC}	99.98 ± 0.040	98.10 ± 0.380	99.72 ± 0.150	99.54 ± 0.190	98.20 ± 0.370
xv_{n_v}	100.00 ± 0.000	98.26 ± 0.360	99.54 ± 0.190	99.44 ± 0.210	98.78 ± 0.300

Table 34: Hit rates ($n = 500$, $\tau = 0.50$), with 95 % confidence intervals (all values multiplied by 100).

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	98.74 ± 0.310	79.64 ± 1.120	90.90 ± 0.800	90.52 ± 0.810	82.16 ± 1.060
xv_1	98.74 ± 0.310	79.64 ± 1.120	90.90 ± 0.800	90.48 ± 0.810	82.20 ± 1.060
xv_{CIC}	96.76 ± 0.490	82.56 ± 1.050	93.68 ± 0.670	93.18 ± 0.700	73.44 ± 1.220
xv_{n_v}	98.54 ± 0.330	81.18 ± 1.080	89.80 ± 0.840	90.84 ± 0.800	79.40 ± 1.120

Table 35: Hit rates ($n = 100$, $\tau = 0.75$), with 95 % confidence intervals (all values multiplied by 100).

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	99.98 ± 0.040	97.14 ± 0.460	99.22 ± 0.240	99.70 ± 0.150	96.86 ± 0.480
xv_1	99.98 ± 0.040	97.14 ± 0.460	99.22 ± 0.240	99.70 ± 0.150	96.86 ± 0.480
xv_{CIC}	99.94 ± 0.070	97.78 ± 0.410	99.58 ± 0.180	99.76 ± 0.140	94.06 ± 0.660
xv_{n_v}	99.98 ± 0.040	97.36 ± 0.440	99.16 ± 0.250	99.70 ± 0.150	96.42 ± 0.520

Table 36: Hit rates ($n = 250$, $\tau = 0.75$), with 95 % confidence intervals (all values multiplied by 100).

IC	Clayton	Gumbel	Joe	Frank	Gaussian
AIC	100.00 ± 0.000	99.86 ± 0.100	100.00 ± 0.000	100.00 ± 0.000	99.90 ± 0.090
xv_1	100.00 ± 0.000	99.86 ± 0.100	100.00 ± 0.000	100.00 ± 0.000	99.90 ± 0.090
xv_{CIC}	100.00 ± 0.000	99.86 ± 0.100	100.00 ± 0.000	100.00 ± 0.000	99.52 ± 0.190
xv_{n_v}	100.00 ± 0.000	99.86 ± 0.100	100.00 ± 0.000	100.00 ± 0.000	99.88 ± 0.100

Table 37: Hit rates ($n = 500$, $\tau = 0.75$), with 95 % confidence intervals (all values multiplied by 100).