A Appendix: Copulas

If we are in the situation where we do know about the marginal distributions but relatively little about the joint distribution, and want to build a simulation model, then the copula is a useful method to deriving such a joint distributions respecting all those given marginal distributions. Here we present some basic ideas of copulas and introduce the two copulas that we used in this paper. For further details on the topic, we refer the readers to[1,2].

Definition 1. *A p-dimensional copula* $C\colon [0,1]^p \to [0,1]$ *is a function which is a cumulative distribution function with uniform marginals.*

The following result due to Sklar (1959) says that one can always express any distribution function in terms of a copula of its margins.

Theorem 1 (Sklar's theorem)**.** *Consider a p-dimensional cdf F with marginals* $F_1,\dots,F_p$. *There exists a copula C, such that*

$$F(x_1,\dots,x_p) = C(F_1(x_1),\dots,F_p(x_p)) \tag{1}$$

for all x_i *in* $[-\infty,\infty]$, $i = 1,\dots,p$. *If* F_i *is continuous for all* $i = 1,\dots,p$ *then C is unique; otherwise C is uniquely determined only on* $\operatorname{Ran} F_1 \times \cdots \times \operatorname{Ran} F_p$, *where* $\operatorname{Ran} F_i$ *denotes the range of the cdf F.*

In our work, we make use of the following two bivariate copulas. The bivariate Gumbel copula or Gumbel-Hougaard copula is given in the following form:

$$C_\gamma(u_1,u_2) = \exp\left\{-[(-\ln u_1)^\gamma + (-\ln u_2)^\gamma]^{\frac{1}{\gamma}}\right\},$$

where $\gamma \in [1,\infty)$. The Clayton copula is given by

$$C_\gamma(u_1,u_2) = \left(\max\{u_1^{-\gamma} + u_2^{-\gamma} - 1, 0\}\right)^{-\frac{1}{\gamma}},$$

where $\gamma \in [-1,\infty) \setminus \{0\}$.

References

1. Hofert, M., Kojadinovic, I., Mächler, M. & Yan, J. *Elements of Copula Modeling with R* (Springer International Publishing, 2018).
2. Joe, H. *Dependence Modeling with Copulas* (Chapman and Hall/CRC, 2014).
3. Minaee, S. *et al.* Image segmentation using deep learning: A survey. *IEEE transactions on pattern analysis machine intelligence* **44**, 3523–3542 (2021).
4. Lintusaari, J. *et al.* Elfi: Engine for likelihood-free inference. *J. Mach. Learn. Res.* **19**, 1–7 (2018).
5. Kosiorowski, D. & Zawadzki, Z. Depthproc an r package for robust exploration of multidimensional economic phenomena. *arXiv preprint arXiv:1408.4542* (2014).
6. Williams, B., Toussaint, M. & Storkey, A. J. Modelling motion primitives and their timing in biologically executed movements. In *Advances in neural information processing systems*, 1609–1616 (2008).
7. Hubert, M., Rousseeuw, P. J. & Segaert, P. Multivariate functional outlier detection. *Stat. Methods & Appl.* **24**, 177–202 (2015).
8. Cerdeira, J. O. *et al.* Revisiting niche fundamentals with tukey depth. *Methods Ecol. Evol.* **9**, 2349–2361 (2018).
9. Chebana, F. & Ouarda, T. B. Depth-based multivariate descriptive statistics with hydrological applications. *J. Geophys. Res. Atmospheres* **116** (2011).
10. Jörnsten, R. Clustering and classification based on the l1 data depth. *J. Multivar. Analysis* **90**, 67–89 (2004).
11. Mosler, K. Depth statistics. In *Robustness and complex data structures*, 17–34 (Springer, 2013).
12. Kim, S., Mun, B. M. & Bae, S. J. Data depth based support vector machines for predicting corporate bankruptcy. *Appl. Intell.* **48**, 791–804 (2018).
13. Hubert, M., Rousseeuw, P. & Segaert, P. Multivariate and functional classification using depth and distance. *Adv. Data Analysis Classif.* **11**, 445–466 (2017).

14. Hammer, H. L., Yazidi, A. & Rue, H. Estimating tukey depth using incremental quantile estimators. *Pattern Recognit.* **122**, 108339 (2022).

15. Kong, L. & Mizera, I. Quantile tomography: using quantiles with multivariate data. *Stat. Sinica* 1589–1610 (2012).

16. Beaumont, M. A. Approximate bayesian computation. *Annu. review statistics its application* **6**, 379–403 (2019).

17. Goodfellow, I. *et al.* Generative adversarial networks. *Commun. ACM* **63**, 139–144 (2020).

18. Casella, G. & Berger, R. L. *Statistical inference* (Cengage Learning, 2021).

19. Aggarwal, A., Mittal, M. & Battineni, G. Generative adversarial network: An overview of theory and applications. *Int. J. Inf. Manag. Data Insights* **1**, 100004 (2021).

20. Hammer, H. L., Yazidi, A., Bratterud, A., Haugerud, H. & Feng, B. A queue model for reliable forecasting of future CPU consumption. *Mob. Networks Appl.* **23**, 840–853 (2017).

21. Jabot, F., Faure, T. & Dumoulin, N. EasyABC: performing efficient approximate bayesian computation sampling schemes using r. *Methods Ecol. Evol.* **4**, 684–687, DOI: 10.1111/2041-210x.12050 (2013).

22. OpenAI. Gpt-4 technical report (2023). 2303.08774.

23. Vaswani, A. *et al.* Attention is all you need. *Adv. neural information processing systems* **30** (2017).

24. Corvi, R., Cozzolino, D., Poggi, G., Nagano, K. & Verdoliva, L. Intriguing properties of synthetic images: from generative adversarial networks to diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 973–982 (2023).

25. Ho, J., Jain, A. & Abbeel, P. Denoising diffusion probabilistic models. *Adv. Neural Inf. Process. Syst.* **33**, 6840–6851 (2020).