

Ontwikkeling van 'n samevattende generatormodel vir kardioïedmodellering in rigtingstatistiek

DM van Wyk, A Bekker, JT Ferreira, M Arashi

Departement Statistiek, Universiteit van Pretoria, Suid-Afrika

Korresponderende outeur: Delene van Wyk **E-pos:** delenevanwyk7@gmail.com

A unifying generator for cardioid modelling in circular statistics: A flexible cardioid-generated class that can capture inherent skewness and multimodality is introduced. Existing extensions of the cardioid distribution will be shown to be special cases of the proposed class of distributions. The flexible behaviour of the class will be graphically demonstrated. Special cases of new and existing models will be compared in simulation studies and data sets.

Rigtingstatistiek dra by tot die ontginning van betekenisvolle resultate in 'n wye reeks navorsingsvelde, insluitend meteorologie en geologie. Anders as in die lineêre wêreld word datapunte uitgedruk as rigtings in twee of meer dimensies wat rondom 'n gekose nulrigting geroteer word. Die aard van rigtingstatistiek eis andersoortige benaderings vir data-analise, teenoor die statistiese gereedskap vir die lineêre ruimte (Mardia 2000).

Om skeefheid en multimodaliteit in sirkulêre data te hanteer, kan 'n generatorfunksie geïmplementeer word. In hierdie artikel word 'n kardioïedgenerator voorgestel met digtheid van die vorm $f(\theta) = c \cdot g(\cos(\theta - \mu))$, waar $g(\cos(\theta - \mu))$ die generatorfunksie en $c = \left(\sum_{k=0}^{\infty} \frac{g^{(k)}(0)}{k!} \int_{-1}^1 x^k \arcsin(x) dx \right)^{-1}$ die normaliseringskonstante is. Benewens bekende verdelings, soos die kardioïed-Von Mises- en gesirkelde Cauchy-verdelings (Mardia 2000), is daar ook aandag aan die eksponensiële kardioïedverdeling (Paula et al. 2021), sowel as die kardioïed-Von Mises- (Mojtaba & Hossein, 2018) en die nuut ontwikkelde kardioïed-t-verdeling geskenk.

Die kardioïed-t-verdeling is gebaseer op werk oor digtheid deur Siew, Kato & Shimizu (2008): $f_{ct}(\theta) = c_{ct} \left(1 + \frac{\kappa}{n} \cos(\theta - \mu) \right)^{-n}$, vir $\kappa > 0$ en n 'n nie-negatiewe heelgetal, waar $c_{ct} = \left[\int_0^{2\pi} \left(1 + \frac{\kappa}{n} \cos(\theta - \mu) \right)^{-n} d\theta \right]^{-1}$.

Statistiese metodes is ingespan om spesiale gevalle van die generatormodel uit te brei, te visualiseer en resultate vanaf data te analiseer. Die maksimum aanneemlikheidsberamers vir die kardioïed-t-verdeling is bepaal, gebaseer op die partiële afgeleides van die log-waarskynlikheid, $\ln L(\theta) = m \ln(c_{ct}) - mn \sum_{i=1}^m \ln \left(1 + \frac{\kappa}{n} \cos(\theta - \mu) \right)$.

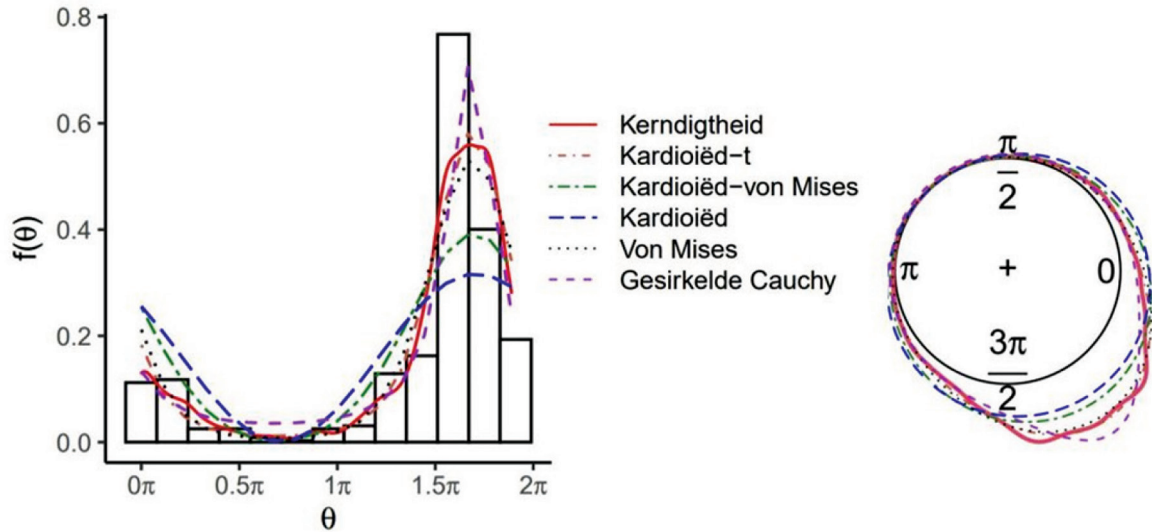
Vanuit simulasiestudies met R-pakkette *LearnBayes* en *gibbs.met* is bevind dat die kardioïed-Von Mises en kardioïed-t-verdelings akkurate maksimum aanneemlikheidsberamings met klein gemiddelde kwadraatfoute oplewer.

Om die doeltreffendheid van die voorgestelde generatormodel te toets, is vyf verdelings (lede van die generatormodel) op die data van verdwyningsrigting van 714 eende gepas en vergelykend geëvalueer deur van Akaike en Bayes se inligtingskriteria (AIC en BIC) gebruik te maak – sien resultate in Tabel I en Figuur 1. Die data, met nul grade as die verwysingsrigting, is oor 'n tydperk van een jaar in Gloucestershire, Engeland, ingesamel (Mardia 1972).

Dit is visueel duidelik dat die kardioïed-t en Von Mises-digthede die nouste verbintenis met die kerndigtheid blyk te hê en dat die digtheidsvoorstellings soortgelyk voorkom in beide die kartesiese (links) en die sirkulêre (regs) geval. Dit word bevestig in Tabel I deur die kleinste AIC- en BIC-waardes vir die kardioïed-t, gevolg deur die kardioïed-Von Mises-verdelings. Dit dui daarop dat die kardioïed-t-verdeling hier die optimale model vir hierdie datastel is.

Tabel I: Maksimum aaneemlikheidsberamings vir lede van die generatormodel.

| Verdeling | $\hat{\alpha}$ | $\hat{\kappa}$ | $\hat{\mu}$ | \hat{n} | AIC | BIC |
|---------------------|----------------|----------------|-------------|-----------|------------------|------------------|
| Kardioïed-t | | 1.8554 | 2.1608 | 5 | 1745.6187 | 1754.7604 |
| Kardioïed-Von Mises | 0.4874 | | 5.3342 | | 1851.0960 | 1860.2370 |
| Kardioïed | 0.9919 | | 5.3706 | | 1967.1060 | 1976.2480 |
| Von Mises | | 2.1161 | 5.3105 | | 1758.9200 | 1768.0620 |
| Gesirkelde Cauchy | 0.6356 | | 5.2859 | | 1782.8420 | 1790.9840 |



Figuur 1: Digheidsvoorstellings van lede van die generatormodel.

Verwysings

Mardia, K.V., 1972, *Statistics of Directional Data*, eerste uitgawe. Academic Press, London.
 Mardia, K.V., Jupp, P.E., 2000, *Directional Statistics*, tweede uitgawe. Wiley, New York. <https://doi.org/10.1002/9780470316979>.
 Mojtaba, H., Hossein, A.M., 2018, Transformation of circular random variables based on circular distribution functions, *Filomat* 32(17), 5931-5947. <https://doi.org/10.2298/FIL1817931M>.
 Paula, F.V., Nascimento, A.D., Amaral, G.J., et al., 2021, Generalized cardioid distributions for circular data analysis, *Stats* 4(3), 634-649. <https://doi.org/10.3390/stats4030038>.
 Siew, H.Y., Kato, S., Shimizu, K., 2008, The generalized t-distribution on the circle. *Japanese Journal of Applied Statistics* 37(1), 1-16. <https://doi.org/10.5023/jappstat.37.1>.