Exercise 2

Load the `tz_malaria` dataset from the RiskMap package and convert it to an `sf` object to handle spatial data.

1. Fit a binomial generalized linear mixed model (GLMM) with a random intercept for cluster.number, assuming no fixed effects. Write down the equations that defines the linear predictor of this model and list the model assumptions; denote this model with M_0. Extract the estimated random effects and use them to compute and plot the empirical variogram. Interpret the variogram to assess spatial correlation.
2. Fit a binomial GLMM that includes temperature and a linear spline transformation for temperature, along with EVI as predictors. Write down the equations that defines the linear predictor of this model and list the model assumptions; denote this model with M_1. Extract the estimated random effects and compute the empirical variogram. Compare this variogram with the one from the null model.
3. Extend both M_0 and M_1 so that they also include a stationary and isotropic spatial Gaussian process with exponential correlation function. Fit the two models using Monte Carlo Maximum Likelihood (MCML) and interpret the summaries of the model fits.
4. Compute the theoretical variograms for both models using their estimated parameters. Overlay the theoretical variograms in a single plot and ensure that the curves are distinguishable with a legend. Discuss whether the inclusion of temperature and EVI reduces spatial correlation in the residuals and how this affects model interpretation.
