Homework assignment 1: Spatial Epidemiology 2025–2026

Deadline: 19/11/2025

This is a group assignments (3 or 4 students per group)

Questions? Ask via Blackboard or mail (christel.faes@uhasselt.be)

Introduction to problem

Colorectal cancer screening is a preventive medical test used to detect early signs of colon or rectal cancer. The goal of this screening is to identify cancerous or precancerous changes in the colon (large intestine) or rectum before symptoms develop, making treatment more effective. Colorectal cancer is one of the most common forms of cancer, and early detection significantly improves survival rates. In Flanders, the program targets adults aged 50 to 74. The screening test is free of charge for the target population, as it is funded by the government.

In this project, we will investigate the spatial distribution of participation to colorectal cancer screening in Flanders.

Description of data

The following data is available:

- The number of invited individuals per gender, age group (5 age-years per group) and municipality.
- The number of individuals participating per gender, age group (5 age-years per group) and municipality.
- Naam = municipality
- Shape-file of Flanders

Note that values in the above dataset in red are imputed values.

Questions

The questions of interest are:

1. Are there areas with a lower participation in the screening as compared to participation in the whole of Flanders? To answer this question, do the following:
   (a) Calculate the number of expected participants per municipality, using the indirect age-standardization with the whole study region (Flanders) as standard population. Calculate this separately for males and females.
   (b) Do some municipalities have a standardized participation ratio (significantly) lower than 1 for males and/or females?
   (c) Use the BYM model to model the total number of participating individuals per area (per gender). Estimate the parameter estimates, relative risks and exceedance probabilities ($P(RR > 1|y)$). What information does the parameter estimates give use? Do any of these relative risks indicate a lower/higher participation?
2. Compare INLA with MCMC: Obtain estimates from the BYM model with both MCMC and INLA. Compare the estimates as well as the computation time.
3. Compare the model with other possible models: Poisson-lognormal, Besag, and BYM2. You can use either MCMC or INLA. Which model is best fitting to the data? Use two types of diagnostics to compare the model fit. Give the results of the best fitting model with your conclusions (if different from the already discussed BYM model).
4. Is the spatial distribution of participation different between males and females? Fit a model on the data of males and females together that allows for (1) a different overall participation rate; (2) a different spatial distribution. Write out the model that you fit. Give the results and your conclusion. (Hint: you will have to fit a model similar to a space-time model!)

In your project, make sure you carefully describe the used model, as well as the model results. A question-answer style report is expected.