DETECTING OUTLIERS IN MORTALITY DATA

A PREPRINT

Matz A. Haugen
matzhaugen@gmail.com

Dorothea Gilbert

March 2, 2022

ABSTRACT

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Keywords Mortality · Outliers

1 Introduction

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

2 Methods

3 Results

References

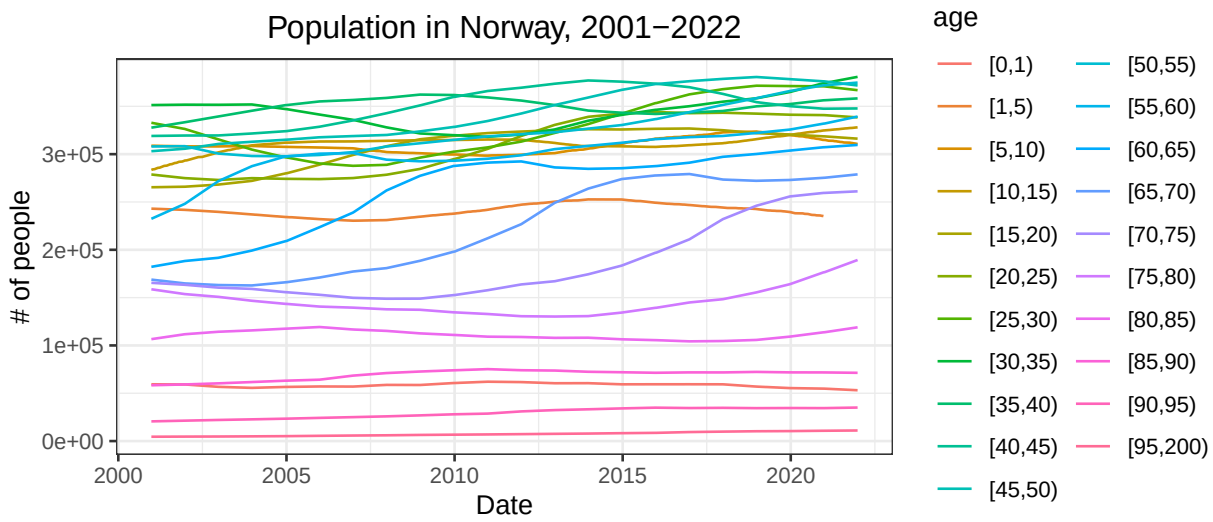

Figure 1: Population in Norway 2001-2022. Data is collected annually at the end of each year and interpolated to a weekly resolution.

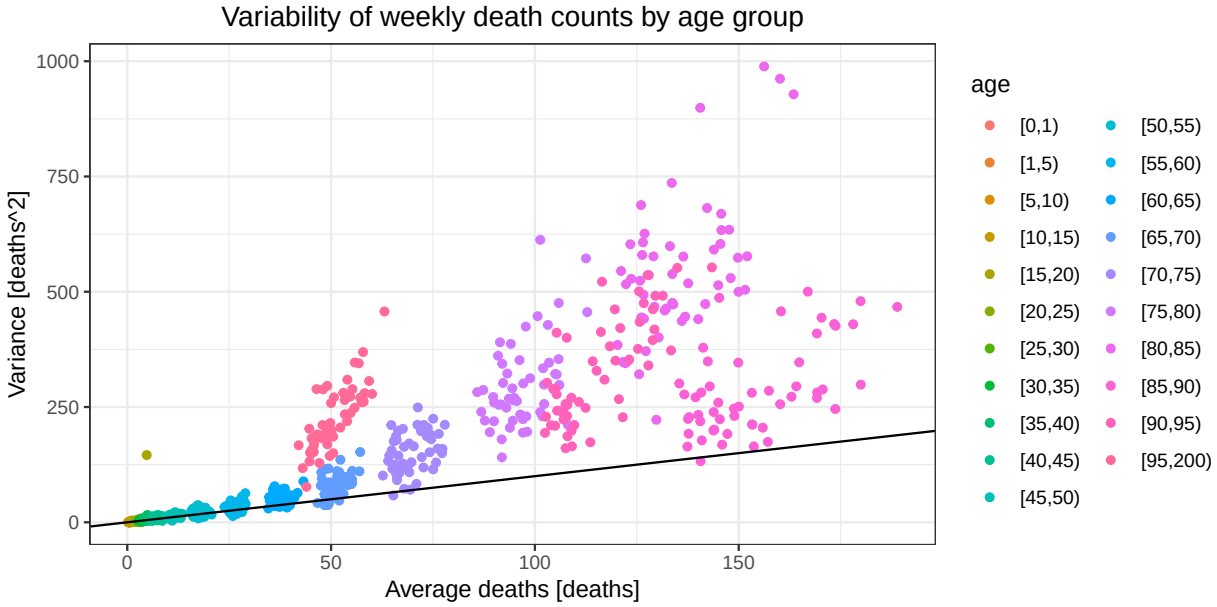

Figure 2: Weekly means and variance for mortality data from 2001-2022, separated into age-brackets of 5 years.

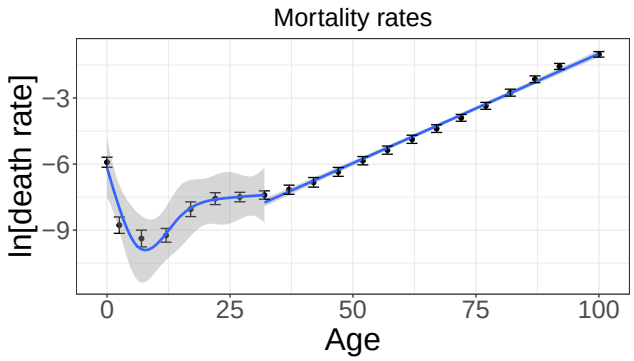

Figure 3: Gompertz law of mortality in Norway: Log-transformed death rates plotted against age using a 5-year age bracket. The average age is taken as the observation on the abscissa. A linear trend is fitted with a weighted squared error loss after age 30 and a natural spline with 4 degrees of freedom before age 30.

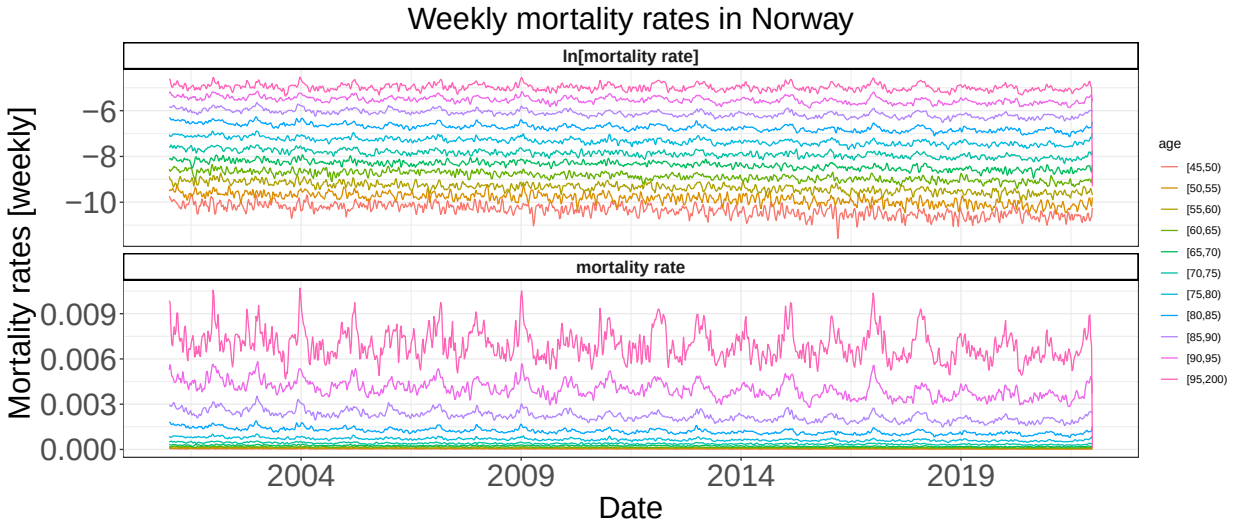

Figure 4: Mortality rates separated into age brackets. Top: Natural-Log-transformed mortality rates. Bottom: Raw mortality rates. Units are different in the y-axis due to the transformation. This is to show both the linear rise in mortality with age and the corresponding change in temporal variance. Notice on the top how variance is lowest in the middle ages.

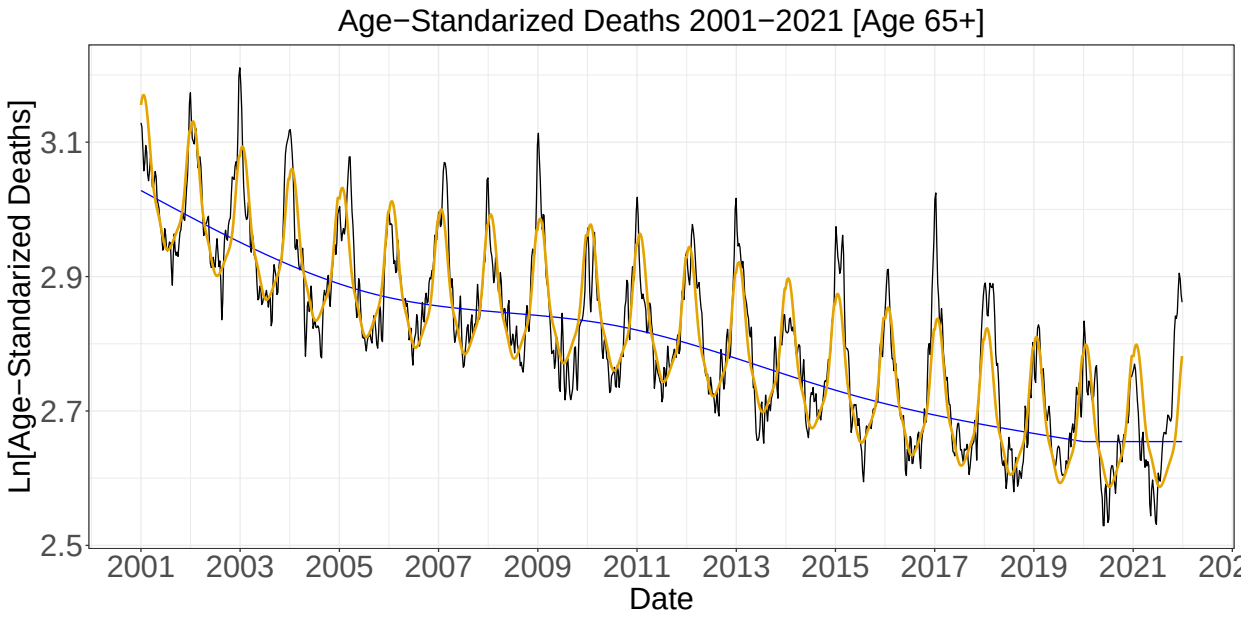

Figure 5: Log-transformed mortality rate for people over 65 years in Norway after age-standardization, using annual population in 5-year age brackets. A trend is added with two splines one for decadal trends (4 degrees of freedom) and one for seasonal trends (7 degrees of freedom). A constant trend is assumed in the years 2020-2021.

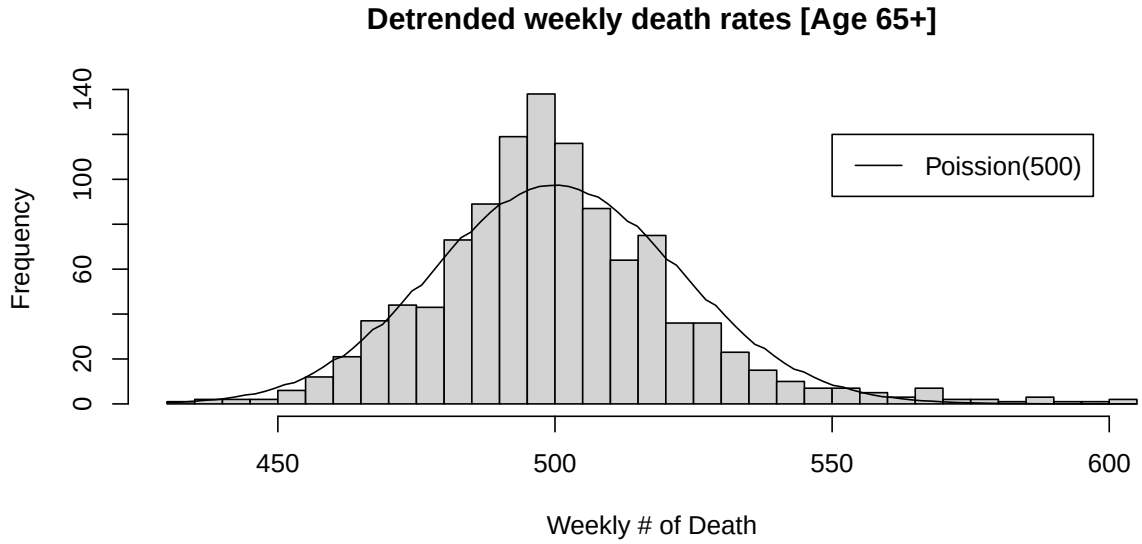

Figure 6: Histogram of weekly deaths for ages above 65 in Norway, juxtaposed with a Poisson distribution with a mean/variance of 500 .

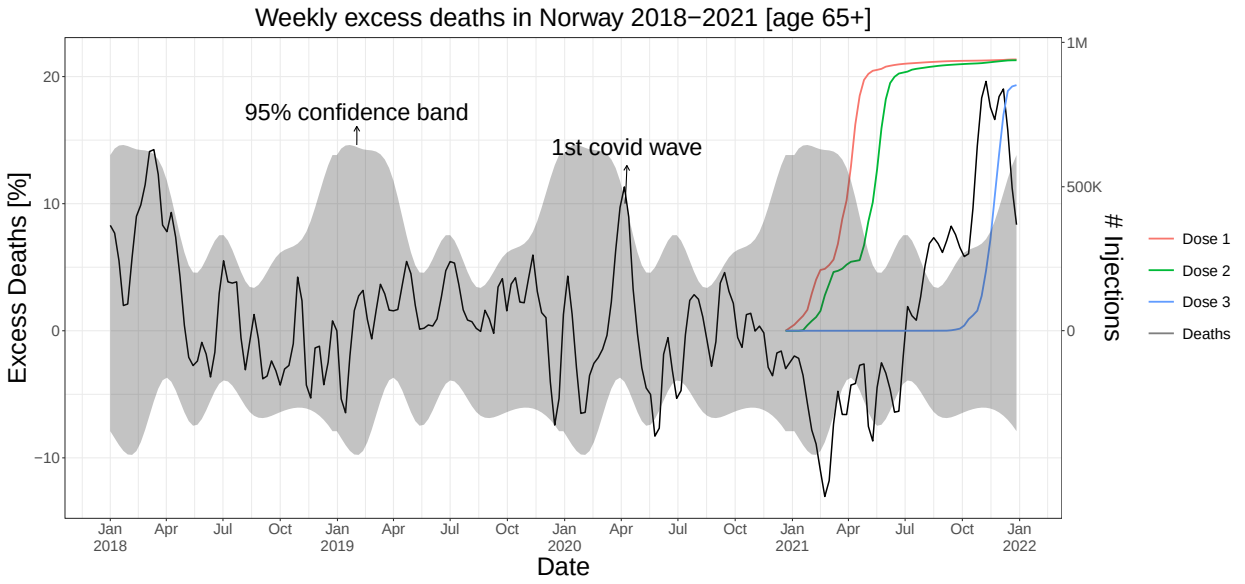

Figure 7: Excess deaths for people over 65 years in Norway after age-standardization and detrending with a seasonal and decadal trend. Trends are fitted robustly at the 50^{th} quantile, while the confidence interval fits the 97.5% and the 2.5% quantiles respectively. Vaccination data is superimposed for the same age group, separated by injection number.

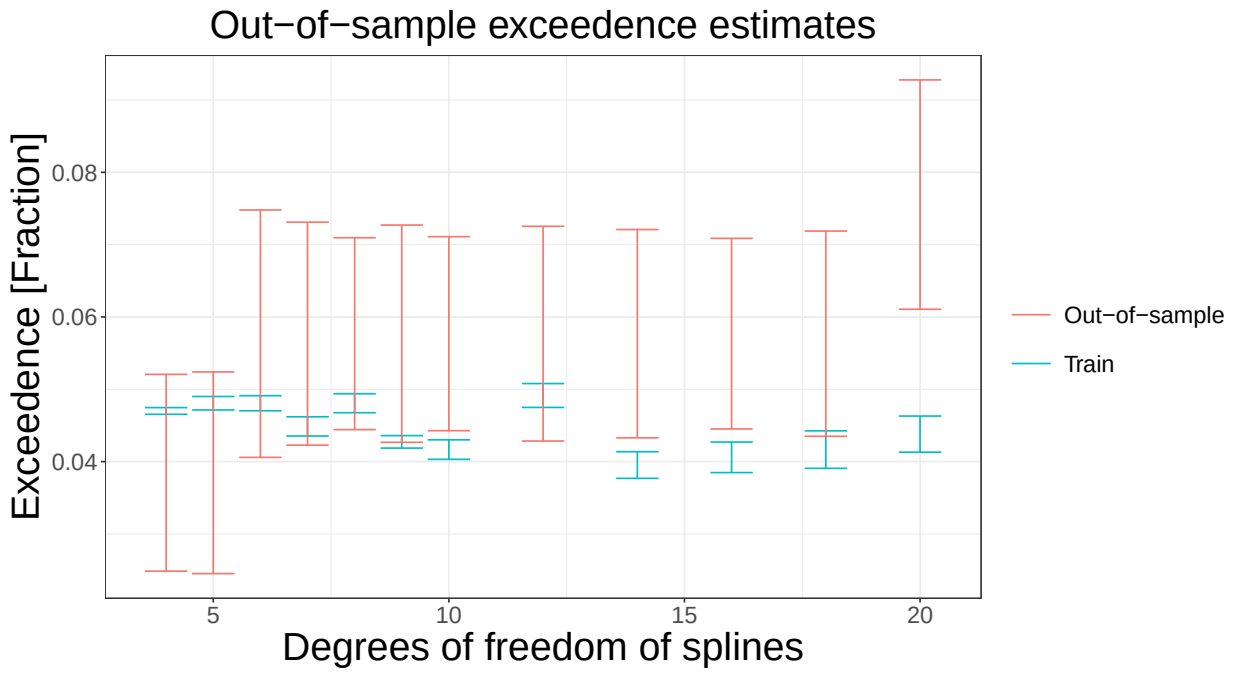

Figure 8: Cross-Validated Exceedence Estimates: An estimate of the exceedence fraction of annual observations from 2001-2019 outside the 95^{th} confidence band for the training period (blue) and for the out-of-sample hold-out year (in red). Each year is held out in turn while the other years are used to compute the confidence band. Error bars are one standard deviation of the average estimate.