Probability, Statistics and Risk, MVE240

Project 4

In the project description we sketch the analysis of the problem we expect you to do. (Obviously you are welcome to do more.) To pass the project a short report should be written and handed in to the project supervisor. In addition the group should present their results in class. The presentation should take about 15 minutes. Please include a short introduction which will facilitate for other students to understand the results of the project. (Do not assume that the audience knows the subject.)

1   Introduction - Poisson Regression

Poisson regression is a tool to handle linear models where the dependent variable counts numbers of events (accidents). One tries to relate the intensity of events to some exploratory variables.

Denote by μ the expected value for a Poisson variate N. Further, let x be an explanatory variable. Traditionally, one uses a Poisson loglinear model of the form

$$\ln \mu = \alpha + \beta x. \tag{1}$$

Intuitively, we can think of the explanatory variable x in Eq. (1) as an index, or a label, related to a certain Poisson process. As we will see in the example below, one often wants to compare the number of occurences for two (or more) situations. Other examples: number of accidents for airline A vs. airline B in one year; number of coronary-disease deaths among smokers vs. non-smokers.

For the model in Eq. (1), the mean satisfies the exponential relationship

$$\mu = \exp(\alpha + \beta x) = \mathrm{e}^{\alpha}(\mathrm{e}^{\beta})^{x}$$

From this formula, one notes that a one-unit increase in x has a multiplicative impact of e^{β} on μ. If $\beta = 0$, then $\mathrm{e}^{\beta} = \mathrm{e}^{0} = 1$ and the multiplicative factor is 1; that is, the mean of N does not change as x changes. If $\beta > 0$, then $\exp(\beta) > 1$, and the mean of N increases as x increases.

Read the Example 7.18 in course book and discuss it in yours report. Perform similar analysis for the following problem.

2   Vehicle Accidents for Elderly Drivers

A study dealt with motor-vehicle accident rates for elderly drivers. The sample consisted of 16262 Medicaid enrolles aged 65-84 years, with data on each subject for a period of somewhere between 0 and 4 years.

The total observation time for women in the sample was 17.3 thousand years. During this period, they had 175 accidents in which an injury occurred. The total observation time for men was 21.4 thousand years, during which they had 320 injurious accidents. From this information, calculate estimates of the intensity for women, λ_{w}, and for men, λ_{m}. Use the unit "crashes per thousand years". Discuss the results.