Simple, flexible modeling for integer value responses with an application to alcohol consumption analysis

Author: Mr Claude Nadeau

Submission ID: 602

Format: CPS Paper

Reference Number: 602

Brief Description

Various ways to model an integer valued response variable N exist (e.g., zero inflated Poisson, Negative Binomial, etc.).

We tackle integer response modeling by simply viewing N as being the integer part of a positive continuous response T.

So observing N=8 is akin to observing that continuous T is between 8 and 9 (interval censoring).

We are then naturally drawn to survival analysis models.

Though several models exists (e.g., Weibull, log-Normal, Cox proportional hazard), we chose to use piecewise constant hazard which is simple, flexible and able to fit adequately.

An R package (ModIvIc) was created to assist in model specification and fitting.

Using data from the Canadian Community Health Survey (CCHS), this approach was utilised to model weekly alcohol consumption as a function of covariates such as smoking and age.

Abstract

Various ways to model an integer valued response variable N exist (e.g., zero inflated Poisson, Negative Binomial, etc.). We tackle integer response modeling by simply viewing N as being the integer part of a positive continuous response T. So observing N=8 is akin to observing that continuous T is between 8 and 9 (interval censoring). We are then naturally drawn to survival analysis models. Though several models exists (e.g., Weibull, log-Normal, Cox proportional hazard), we chose to use piecewise constant hazard which is simple, flexible and able to fit adequately. An R package (ModIvIc) was created to assist in model specification and fitting. Using data from the Canadian Community Health Survey (CCHS), this approach was utilised to model weekly alcohol consumption as a function of covariates such as smoking and age.