Assignment 9 - Poisson Regression. Due November 23, 11:59pm 2018

EPIB607 - Inferential Statistics[a]

[a]Fall 2018, McGill University

This version was compiled on November 15, 2018

In this assignment you will practice Poisson regression. Be sure to state the regression model in terms of parameters first. Use regression functions to fit these models. Answers should be given in full sentences (DO NOT just provide the number). All figures should have appropriately labeled axes, titles and captions (if necessary). Units for means and CIs should be provided. All graphs and calculations are to be completed in an R Markdown document. Please submit both the compiled HTML report and the source file (.Rmd) to myCourses by November 23, 2018, 11:59pm. Both HTML and .Rmd files should be saved as 'IDnumber_LastName_FirstName_EPIB607_A9'.

Poisson regression | Rate ratio | Rate difference | Person time | Offset

Template

There is no template for this assignment. You may use the same template from previous assignments. Be sure to include your code at the end of the compiled report.

1. Bednets

See the 2018 Lancet article *Efficacy of Olyset Duo, a bednet containing pyriproxyfen and permethrin, versus a permethrin-only net against clinical malaria in an area with highly pyrethroid-resistant vectors in rural Burkina Faso: a cluster-randomised controlled trial* (`Bednets.pdf` in `A9` folder of myCourses) by Tiono et. al.

a. Reproduce the Rate ratio and 95% CI in Table 2. Start by writing down the model in terms of a regression equation with parameters and determinants. Interpret the rate ratio.
b. Perform a goodness of fit test for the model in part (a). Is this a good fit? If not, explain what you might do to improve the fit.
c. Calculate the rate difference and 95% CI comparing PPF-treated to Standard long-lasting insecticidal nets.

2. Population mortality rates in Denmark

The following table contains mortality data for males and females in Denmark for 4 age groups over three time periods.

Year	Age	Female_deaths	Female_PT	Female_rate	Male_deaths	Male_PT	Male_rate
1980-1984	70-74	15989	586882.8	0.0272439	23810	456908.21	0.0521111
1980-1984	75-79	20838	454142.7	0.0458843	24707	300318.92	0.0822692
1980-1984	80-84	24073	297678.6	0.0808691	20319	167303.51	0.1214499
1980-1984	85-89	20216	147771.7	0.1368057	13524	74295.83	0.1820291
2000-2004	70-74	13912	521561.9	0.0266737	17360	436994.92	0.0397259
2000-2004	75-79	19731	471945.5	0.0418078	22477	341362.82	0.0658449
2000-2004	80-84	25541	369989.9	0.0690316	22992	217929.72	0.1055019
2000-2004	85-89	27135	226798.1	0.1196439	17444	104009.58	0.1677153
2005-2009	70-74	12179	540568.6	0.0225300	15782	472012.84	0.0334355
2005-2009	75-79	17273	444474.2	0.0388616	19547	344351.34	0.0567647
2005-2009	80-84	23513	363534.1	0.0646789	21781	230530.24	0.0944822
2005-2009	85-89	26842	237877.3	0.1128397	17811	114485.04	0.1555749

a. Come up with a suitable regression model for this data. Write down the regression equation in terms of parameters and determinants.
b. Estimate the parameters of this model using the data in the table above. Provide the fitted regression equation.
c. Interpret the parameter for gender. Are mortality rates significantly different in males compared with females?
d. Perform a goodness of fit test for the fitted model in part (b). Is this a good fit?