Variables influencing the number of members in a household domiciled in SOCCSKSARGEN, a region of the Philippines

Author: Urszula Wolanowska, School of Mathematics & Statistics, University of Glasgow, Glasgow
Email: 2322312w@student.gla.ac.uk

Introduction

- It is of interest to the government to investigate variables that affect the number of people living in one household in the Philippines.
- Data on 2122 households is obtained from the **Family Income and Expenditure Survey (FIES).**
- My analysis deals with the SOCCSKSARGEN region of the Philippines (Figure 1).

Figure 1: Map of the Philippines with the SOCCSKSARGEN region highlighted in red

Variables from **FIES**:
Response (y): *Number of Household Members*
Explanatory variables (X): *Household Income* (in Philippine peso ₱), *Food Expenditure (in* ₱), *Head' of Household Gender*, *Head' of Household Age*, *Type of Household*, *House Floor Area, House Age, Number of Bedrooms, Electricity Availability*

Approach and Methodology

- From Figure 2, it can be seen that the number of people living in one household, in other words our response variable y, follows a Poisson distribution such that: $y \sim Po(\mu)$.
- Here, the parameter μ corresponds to the average number of people living in a household.
- This is equivalent to saying that the probability distribution function of y is: $f(y) = \frac{\mu^y e^{-\mu}}{y!}$.

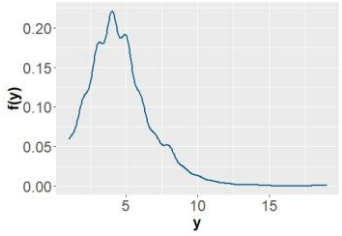

Figure 2: Kernel estimation of the density function of the total number of household members (y)

The relationship between the number of people living in a household $\boldsymbol{y}$ and the predictor variables **X** can be modelled using Poisson, Quasi-Poisson and Negative Binomial Regression.

Results

The best model fit was achieved with the **Quasi-Poisson regression with log-transformed covariates**:

$$E(y) = log\ \hat{\mu} = \log(\boldsymbol{x}^T)\widehat{\boldsymbol{\beta}}, \quad (1)$$

For Quasi-Poisson $Var(y) = \phi\hat{\mu}$ where $\phi = \frac{Pearson's\ chi-squared\ statistics}{Residual\ degrees\ of\ freedom} = \frac{X^2}{n-p}$ is the dispersion parameter for this model.

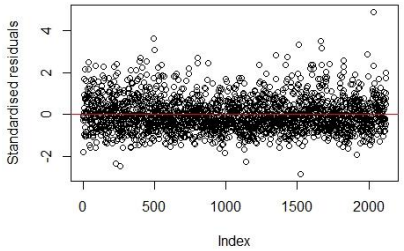

Figure 3: Standardised Residuals plot for model (1)

- Parameter estimates $\widehat{\boldsymbol{\beta}}$ are all significant.**There is a significant relationship between all the explanatory variables and the response.**
- Nonetheless, there is still some variation unexplained by this model. Shown in Figure 3 where plenty of standardised residuals are above 2.

Conclusion and discussion

- The explanatory variables provided have a significant impact on the Number of Household Members.
- However, **more predictors** could be included to account for more variation in the response variable. Potential covariates could include Head's of Household Marital Status, Employment Status, interactions and many others.

References

Figure 1, PhilAtlas, Map data © OpenStreetMap contributors. *2020. SOCCSKSARGEN (Region XII). [Online]. [Accessed 29 October 2020]. Available from:* https://www.philatlas.com/mindanao/r12.html