Insight into property prices in Bournemouth County, England

Davide Perez
University of Glasgow, School of Mathematics and Statistics

Introduction

Estate agency companies are often keen on improving their field knowledge about the market and what features of the houses are most valuable in their worth.
Bournemouth County area is of particular interest for our research on some of the most prominent factors in the evaluation of the property price such as:

1. Variation of the cost with respect to the size of the building;
2. Effect of energy efficiency;
3. Most and least affordable regions.

Data and Methodology

The data set has been provided by researchers at University College London, comprising information from the official open Land Registry PPD with structural details and energy performance (guaranteed by EPCs).
From these, a GLM was developed for *price* response explained by the following coviariates:

- *Total floor area*
- *Number of habitable rooms*
- *Type of the properties* (*classified as* Maisonette, Flat, House, Bungalow)
- *Current energy efficiency (based on cost of energy)*
- *Environmental Impact Rating (based on CO2 emissions)*

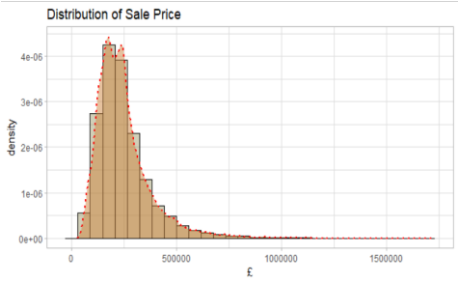

Figure 1: Histogram and density of *price* variable

Having a look on the distribution it is clear how the response variable is right-skewed, therefore the most suitable model to adjust for the assumptions is a Gamma distribution with a log-link function:

$$E[\log(price)] = X^T\beta$$

Fitting of the model

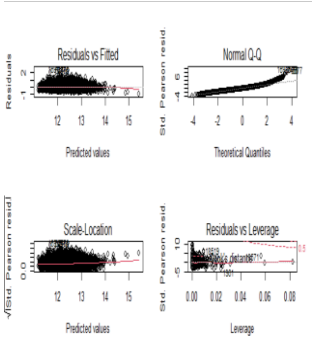

Figure 2 : Check of the diagnostic plots for fitted model

The diagnostic confirms the suitability of the model:

- Residuals vs Fitted looks goos so linearity of the predictors with respect to the log(price) is reasonable
- In the QQ plot the shape of the data cloud on the right confirms the right-skewness
- Scale-Location do not suggest any particular behaviour of the residuals
- Some regression outliers could be present but they actually do not affect significantly the estimates

Results

Table1 allows us to comment on the model insights:

1. **Relationship size-price:** Since flats are the reference baseline for the type of property, keeping all the variables fixed the price will increment of the 1% with respect to a unitary increase on the total floor area
2. **Energy efficiency:** The coefficients of interest suggest that on average expensive houses are also more efficient and heve less carbon dioxide emissions

term	estimate	std.error	statistic	p.value
(Intercept)	10.7773025	0.0123115	875.385064	0.0000000
tfarea	0.0133869	0.0001289	103.884588	0.0000000
numberrooms	0.1087063	0.0027948	38.896106	0.0000000
propertytype_epcHouse	0.3213887	0.0138774	23.159127	0.0000000
propertytype_epcMaisonette	0.1655526	0.0277583	5.964086	0.0000000
propertytype_epcBungalow	0.6365825	0.0170972	37.233162	0.0000000
CURRENT_ENERGY_EFFICIENCY	0.0012294	0.0003328	3.693999	0.0002212
ENVIRONMENT_IMPACT_CURRENT	0.0017427	0.0003070	5.676807	0.0000000
tfarea:numberrooms	-0.0006595	0.0000199	-33.187726	0.0000000
tfarea:propertytype_epcHouse	-0.0029121	0.0001554	-18.739541	0.0000000
tfarea:propertytype_epcMaisonette	-0.0026165	0.0003478	-7.523360	0.0000000
tfarea:propertytype_epcBungalow	-0.0044207	0.0001975	-22.379607	0.0000000

Table 1: Summary of the fitted Gamma Model on R

3. **Affordability:** Dournemouth data set shows a high concentration of units in the eponymous town. Differences could only be discussed among the East and west constituency; in thr East Part there's more variability and less central values which suggest properties are more expensive

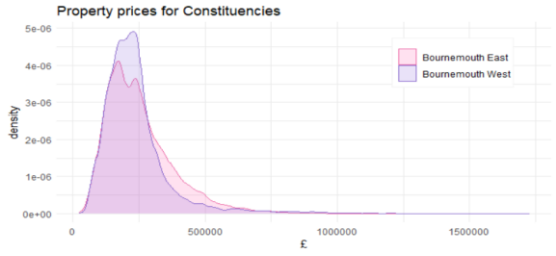

Figure 3 : How prices varies among the constituency