

Hw2
Advance Business Analysis

Total 45 points
Due: February 23

1. SmartAttire, is an established clothing catalog company with an online presence at SmartAttire.com. In order to gain an understanding of the aggregate visiting patterns of broader Internet users, the firm purchased Media Metrix panel data. For a sample of 2728 people who visited any online apparel site at least once during the second half of 2017, the dataset reports how many visits each person made to the SmartAttire.com web site, along with some demographic information. Management would like to know whether the frequency of visits to the web site is related to demographic characteristics.
Since you are an expert, the manager comes to you to help you understand the role of demographic variables in website visits.

[i] First aggregate the data at total visit level and count the number of costumers. Then plot a bar chart. Show your plot and comment on what the distribution looks like. [5]

[ii] Write down a Poisson model that estimates the impact of covariates ((income, sex, age, and size) on number of visits. Show the equation, estimate GLM model, show the output and interpret the estimates. [10]

[iii] You believe that income will go up by 10 percent in the coming year. How will it change the number of visits? We aggregated the data at visit level in part (i). Calculate the updated number of visits and plot the bar chart. Comment on what you see. [15]

2. I am attaching a call center dataset on the number of calls of a call center from people with different income and gender. You see lots of zeros and believe that even though number of calls may follow Poisson model, this is a zero inflated Poisson model (ZIP).

Explain the mechanism behind zero inflated model. What modification to the Poisson model should be made? Explain clearly. [5]

Estimate the models:

1. first (model 1), use log of income as a covariate for zero inflation part and the other two as covariates for Poisson model,
2. and second (model 2), use Gender as covariate for zero inflation part and the other two for Poisson model. [5]

Can you explain what the economic meaning of these covariates is? Plot the actual and predicted values, and the show the deviance. Which model (model 1 or 2) seems like a better fit? [5]