Quote-rate trending analysis

Ranaji Krishna

September 22, 2015

Problem Statement

- To identify whether service providers are becoming more or less inclined to quote over time?

- To explore if there is evidence that product changes over the last two months have caused site-wide shifts in quoting behavior?

Approach

- It was envisaged that site-wide shifts would affect the quoting behavior across all service providers. Therefore quote-rates for invited service providers were computed every hour, and an equally weighted average of the ratios was computed for the hour.

- An estimation model of quote-rate (dependent variable) against time (independent variable) was developed; the significance of time in estimating quote-rate was investigated.

Data Cleaning

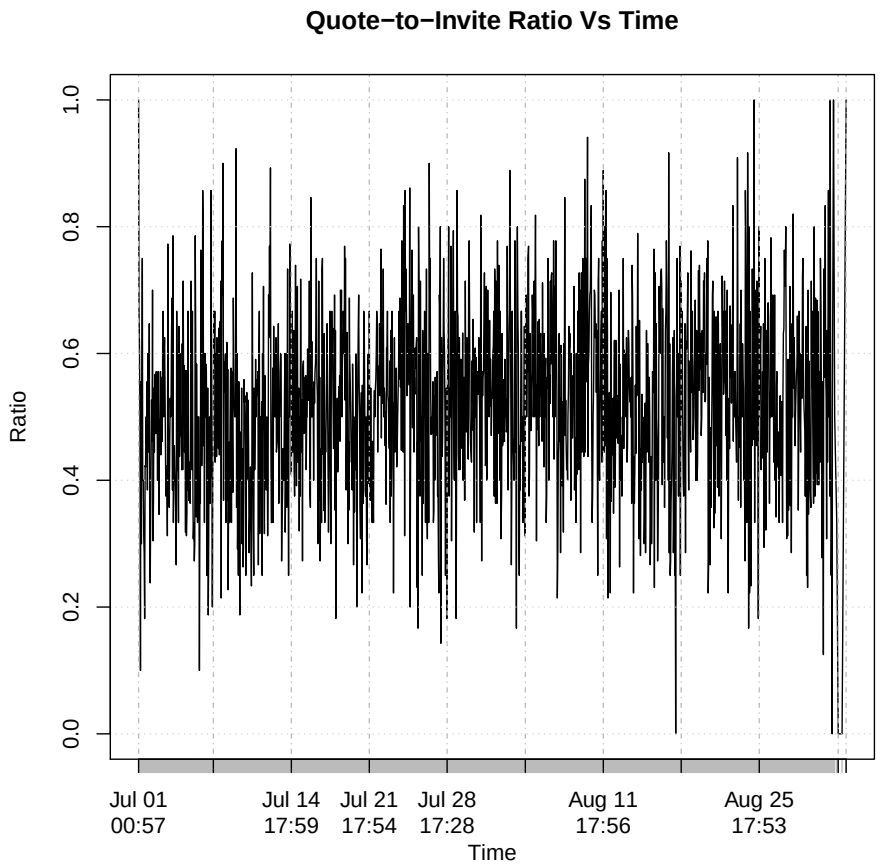
- A table containing entries of "requests sent", "invites sent", "quotes received", "locations", "categories" and "users" was created from relational databases containing historical data using SQLite.

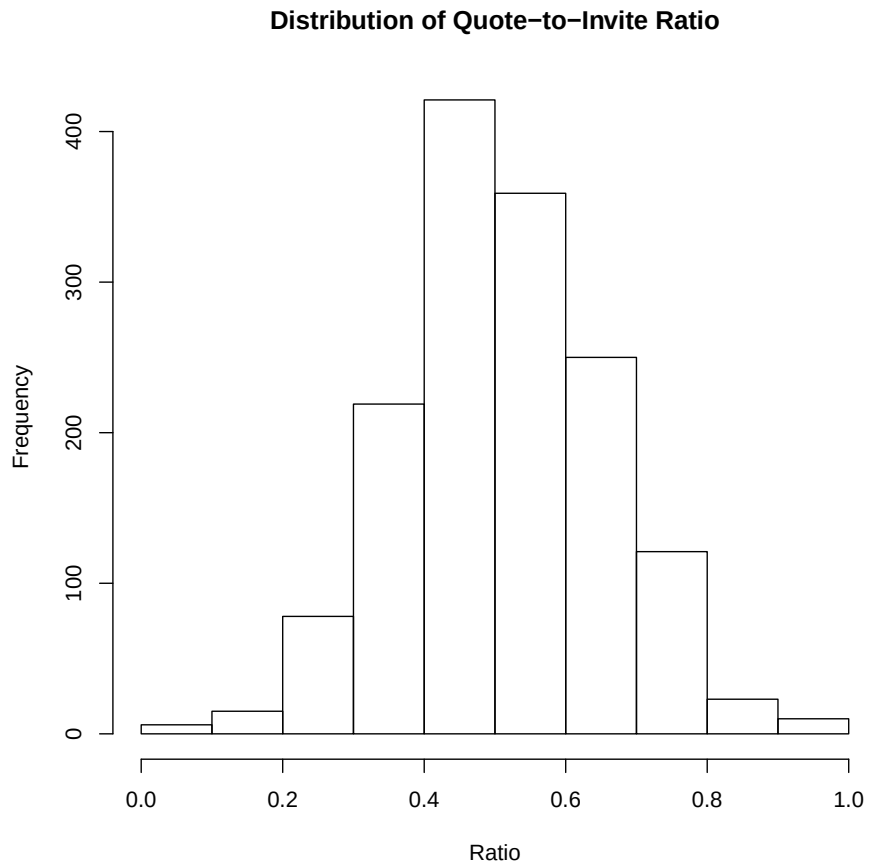
- An average quote-invite ratio was computed for every hour. This was done as follows:
  - for each invited service provider in an hour, a quote-invite ratio was computed by diving the number of invites that led to quotes by the total number of invites received;
  - then an average quote-invite ratio was computed across the service providers for the hour.

- This yielded 1502 quote-invite ratios from the historical data set which contained dates of "invites sent" spanning from 1 July 2013 - 2 September 2013.

Estimation technique

- Trends of quote-invite ratio are shown below:





Estimation technique

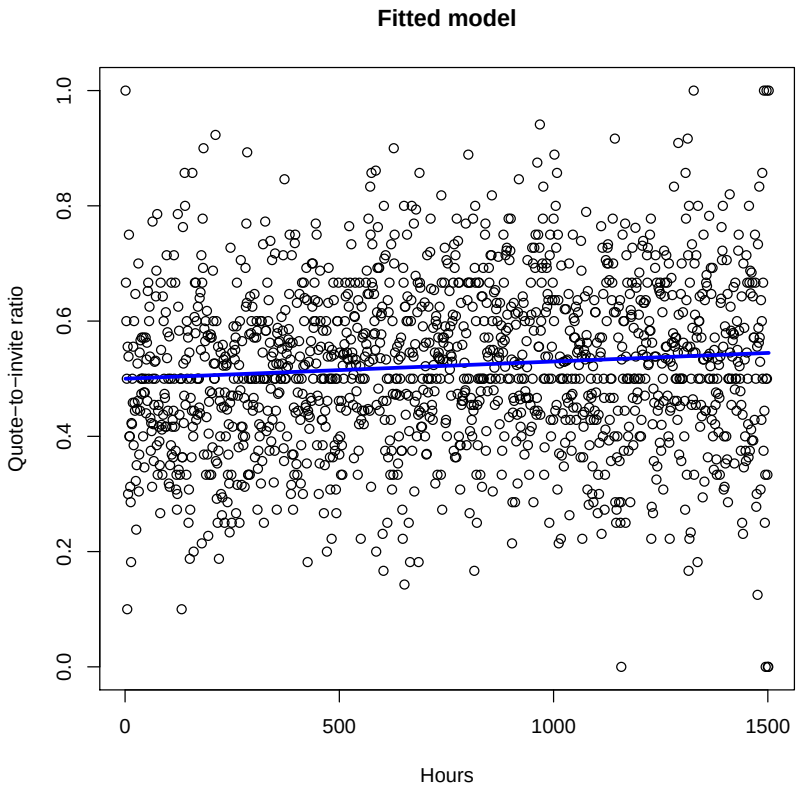
contd.

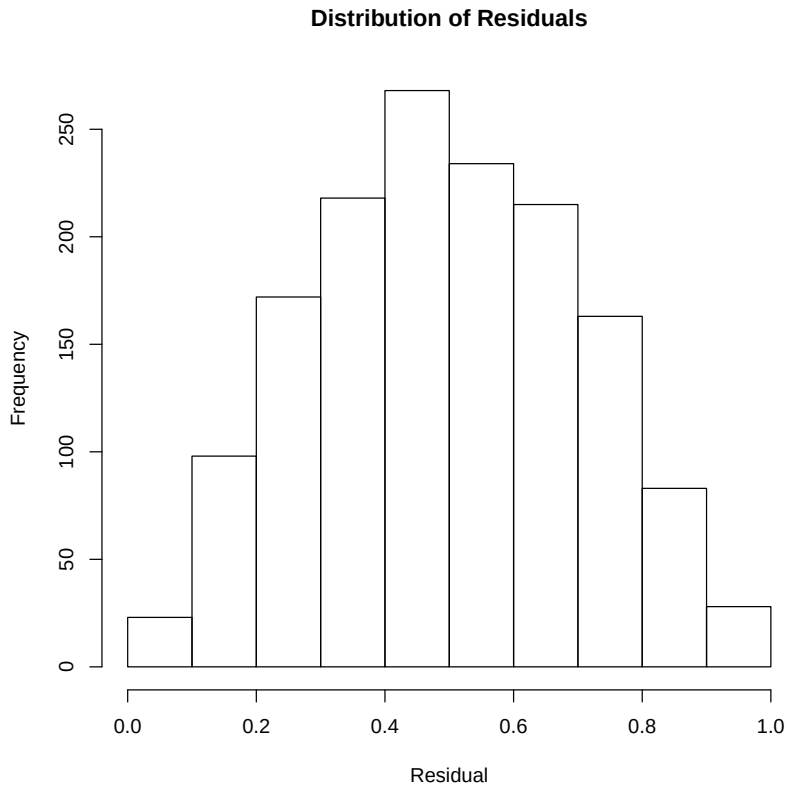
- Regression models involving data from the unit interval, such as proportions, are typically modeled using beta distributions.
- However, beta distribution for the response variable cannot be used due to inclusion of zeros and ones.
- A logit transformation works if we make all zeros slightly more than zero and all ones slightly less than one. However, this introduces a certain amount of arbitrary bias into the model.
- We use a zero-and-one inflated beta distribution model (mixture of a Beta and a Bernoulli distribution) for our purpose.

Results

Coefficient of regression

Coefficient	Estimate	Std. Error	p-value	95% Confidence Intervals
Hour	1.195E-04	2.149E-05	3.17E-08	(7.74e-05, 1.612e-04)





- Computations for fitting the model were carried out using the package **gamlss** in R.

Interpretation of Coefficient

- We note that the "Hour" has significant explanatory power in estimating quote-invite ratio.
- The model can be expressed as:

$$\ln\left(\frac{q/i}{1-q/i}\right) = 1.195e^{-04} \times hr \tag{1}$$

where q/i is the quote-to-invite ratio and hr is the hour variable.

- The coefficient can be interpreted as the percentage change in "quote-invite" to "no quote-invite" ratio, due to 1% increase in hour.

Interpretation of Coefficient

contd.

- A positive coefficient shows improvement in the invite-quote ratio; it reveals that service providers are becoming more inclined to quote over time.

- Example:
  An hour has an average of 9 quotes sent for 10 invites received across invited service providers, generating a "quote-invite" to "no quote-invite" ratio of 9. After 100 days the ratio is expected to increase by 28.68% to 11.58. Therefore after 100 days, for 10 invites sent, 9.21 quotes are expected to be received.

Behavior across Categories

- The analysis was performed for each Category.
- Categories that had coefficients with significant explanatory power were analyzed.
- Categories with positive coefficients:

Categories in which service providers are more inclined to quote

Category	Coefficient	Std. Error	p-value
Home Staging	0.00426	0.00113	0.00024
Wedding Planning	0.00253	0.00058	0.00002
Electrical	0.00204	0.00080	0.01133
Video Filming	0.00182	0.00076	0.01748
Balloon Artistry	0.00114	0.00027	0.00003
Bartending	0.00105	0.00036	0.00376
Tree and Shrub Service	0.00096	0.00031	0.00194

Behavior across Categories

contd.

- Categories with negative coefficients:

Categories in which service providers are less inclined to quote

Category	Coefficient	Std. Error	p-value
Algebra Tutoring	-0.00052	0.00024	0.03117
Dance Lessons	-0.00061	0.00031	0.04709
Tree Trimming	-0.00114	0.00039	0.00341
Concrete Services	-0.00179	0.00060	0.00306
Mold Inspection	-0.00237	0.00079	0.00317
Event Photography	-0.00246	0.00083	0.00330
Local Moving (under 50 miles)	-0.00317	0.00108	0.00400
Roof Repair	-0.00905	0.00402	0.02943
Event Decorator and Designing	-0.01772	0.00809	0.03521

Behavior across Locations

- The analysis was performed for each Location.
- Locations that had coefficients with significant explanatory power were analyzed.
- Locations with positive coefficients:

Locations where service providers are more inclined to quote

Location	Coefficient	Std. Error	p-value
Provo-Orem, UT	0.01226	0.00402	0.00423
Raleigh, NC	0.01147	0.00480	0.02067
Des Moines-West Des Moines, IA	0.00773	0.00437	0.08825
San Jose-Sunnyvale-Santa Clara, CA	0.00289	0.00169	0.09033
San Antonio-New Braunfels, TX	0.00244	0.00146	0.09665

Behavior across Locations

contd.

- Locations with negative coefficients:

Locations where service providers are less inclined to quote

Location	Coefficient	Std. Error	p-value
Deltona-Daytona Beach-Ormond Beach, FL	-0.01145	0.00622	0.07969
Greenville-Anderson-Mauldin, SC	-0.01308	0.00633	0.04639
Stockton-Lodi, CA	-0.01312	0.00515	0.01607
Colorado Springs, CO	-0.03303	0.01629	0.05027

END