**Yaotian Zou**

zouyaotian@hotmail.com

979-739-3681

Avenel, New Jersey, 07001

**SUMMARY**

* 2+ years of data science / predictive modeling experience in auto insurance industry (Pricing, marketing, underwriting, etc.)
* 6+ years advanced statistical modeling techniques (particularly generalized linear models (GLM)).
* 7+ years hands-on advanced SAS programming (extensive use of SAS macro, array and SQL procedures). Proficient in other statistical software such as R and Python.
* 6+ years of academic research in engineering application of advanced statistical methods.
* Current with machine learning techniques such as classification and regression trees (CART), random forests, support vector machine, generalized boosted models (GBM) and clustering methods.

**EXPERIENCE**

**Data Scientist / Predictive Modeler,** 11/2015 – Present

Plymouth Rock Management Company of New Jersey, Red Bank, New Jersey,

* Developed a ground-up vehicle rating model for pricing the personal injury protection (PIP) insurance coverage based on vehicle characteristics and safety features. Improved the premium segmentation by 10% as a pioneer project in auto insurance industry.
* Built marketing response models for direct mail and online leads based on third party demographics and credit data.
* Proposed a customer retention model which allows predicting customers’ life time value at different point of time. Substantially optimized marketing expense strategies in new business acquisition.
* Developed a proprietary model based on insureds’ prior carrier data (e.g. length of insurance, lapse in coverage, etc.) which helped both pricing enhancement and underwriting.
* Developed an unconstrained model for pricing and presented its performance measures against current pricing model to executive management team on a quarterly basis. Quantified the impact of regulation constraints on segmentation and accuracy in premium pricing and tracked its change over time. Automated the process from data preparation to results evaluation which reduced the work load by 75%.
* Provided technical support on model implementation, maintenance and updates.

**Postdoctoral Research Assistant,** 01/2015 – 10/2015

**Graduate Research Assistant (Ph.D.)**, 08/2012 – 12/2014

Center for Road Safety, Purdue University,

* Performed advanced statistical and econometric analysis on motor vehicle crashes to derive insights on the safety performance (e.g. injury risk reduction) of road barriers on behalf of Indiana Department of Transportation (INDOT).
* Assessed the change in injury risks and crash frequencies due to the use of barriers using generalized linear regression models (e.g. logistic regressions and negative binomial regressions), and provided INDOT with strategic alternatives and numerous scenarios for optimal use of different types of barriers, and the results were published in a peer-reviewed journal.
* Provided valuable insight by leveraging multiple analytical and statistical techniques including exploratory data analysis (EDA), linear and non-linear regressions, decision trees, clustering methods, statistical simulation, machine learning and predictive modelling methods.
* Managed the full cycle of analysis including writing proposal, collecting and processing data, programming, analytical and statistical modelling, and presenting the results.

**Graduate Research Assistant (M.S.)**, 08/2010 – 07/2012

Civil Engineering, Texas A&M University,

* Developed a new count model based on the double Poisson distribution and examined its potential application in modelling the motor vehicle crash frequency.
* Programmed its parameter estimation algorithm based on both maximum likelihood estimation and Bayesian methods (Markov Chain Monte Carlo) using R software. The results were presented in a national conference and published in a peer-reviewed journal.

**EDUCATION**

**Ph.D.** in Civil Engineering, Purdue University, West Lafayette, Indiana (GPA: 3.97/4.0) 12/2014

**M.S.** in Civil Engineering**,** Texas A&M University,College Station, Texas (GPA: 4.0/4.0) 08/2012

**B.E.** in Transportation Engineering**,** Southeast University, Nanjing, China (GPA: 87/100) 06/2010

**AWARDS**

2017 President’s Award of Plymouth Rock (top 5% employee based on annual performance review)

**CERTIFICATIONS**

“SAS Certified Base Programmer for SAS 9”, SAS Institute 06/2015 - present

**RELEVANT COURSES**

Statistical Analysis/ Mathematical Statistics/ Applied Categorical Data Analysis/ Modern Applied Statistics/ Bayesian Applied Decision Theory/ Recent Development in Bayesian Analysis/ Statistical and Econometric Methods/ Microeconomics

**SKILLS**

SAS/ SQL/ R/ Python/ JMP/ WINBUGS/ LIMDEP/ ArcGIS/ MS Access/ MS Excel/ C++/ AutoCAD/ Data Mining/ Data Cleaning/ Data Management/ Data Visualization/ Statistical and Econometric Modeling/ Predictive Modeling/ Analysis of Variance/ Regression Analysis/ Machine Learning/ Generalized Linear Model/ Bayesian Estimation/ Survival Analysis/ Time Series/ Transportation Engineering/ Traffic Safety/ Crash Data Analysis

**PUBLICATIONS**

**Zou, Y.** andTarko, A. P. (2018). Barrier-relevant crash modification factors and average costs of crashes on arterial roads in Indiana. *Accident Analysis & Prevention*, Vol. 111, pp. 71-85.

**Zou, Y.** andTarko, A. P. (2016). An insight into the performance of road barriers - redistribution of barrier-relevant crashes. *Accident Analysis & Prevention*, Vol. 96, pp. 152-161.

**Zou, Y.,** Tarko, A. P., Chen, E., and Romero, M. A. (2014). Effectiveness of cable barriers, guardrails, and concrete barrier walls in reducing the risk of injury. *Accident Analysis & Prevention*, Vol. 72, pp. 55-65.

**Zou, Y.,** Geedipally, S.R. and Lord, D. (2013). Evaluating the Double Poisson Generalized Linear Model. *Accident Analysis & Prevention*, Vol. 59, pp. 497-505.

**CONFERENCE PROCEEDINGS**

**Zou, Y.,** Geedipally, S.R. and Lord, D. (2013). Evaluating the Double Poisson Generalized Linear Model. Presented at the 92nd Annual Meeting of the Transportation Research Board, Washington. D.C..

**Zou, Y.**, Lord, D. and Geedipally S.R. (2012). Over- and Under-Dispersed Count Data: Comparing the Conway-Maxwell-Poisson and Double-Poisson Distributions. Presented at the 91st Annual Meeting of the Transportation Research Board, Washington. D.C..