Associations between patient system characteristics and out-of-hospital transport time

Student Researcher: *Austin Holmes*

Faculty Mentor: *Dr. Mohammad Aziz*

Department of Mathematics | University of Wisconsin-Eau Claire

Introduction

The main goal of our project is to find significant patient and system characteristics related to the response variable, out-of-hospital transport time. Research was done in the past using multiple linear regression (MLR) and quantile regression (QR) for evaluating these associations, but insights into skew-normal regression (SNR) were not included.

Data

Out-of-hospital data between the years 2013 and 2016 was received from the Eau Claire Fire Department about patients who were transported via ambulance to the hospital because of either cardiac arrest, chest pain, or trauma.

Dependent Variable:

- Length of Time

Independent Variables:

- Type of Pain
- Zone
- Destination
- Urgency
- Lights/Sirens
- Age
- Gender
- Time of Day

Goals

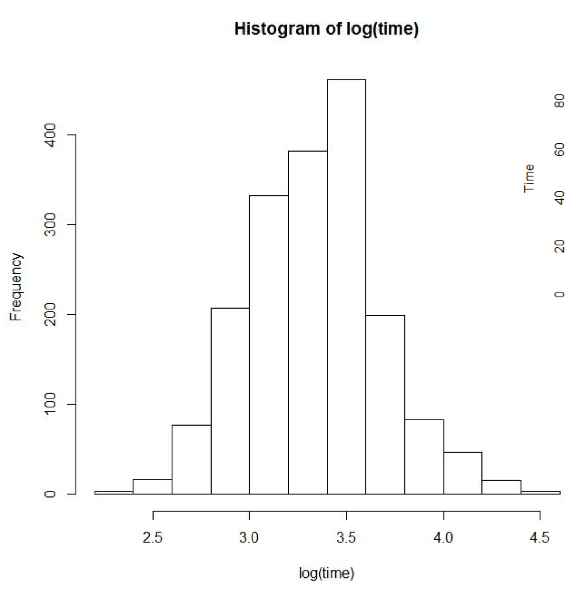
Our primary goals for this project were to find significant predictor(s) related to the response variable and to determine which regression has the better performance in measuring how well out-of-hospital systems function. The three regressions we will be focusing on are:

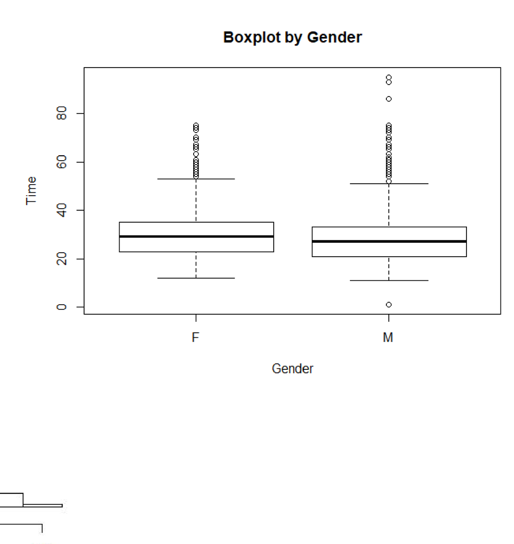
- Multiple Linear Regression (MLR)
- Quantile Regression (QR)
- Skew-Normal Regression (SNR)

Indicators for deciding how well out-of-hospital systems perform have distributions that are skewed, so we hypothesize that an SNR model will be the best fit.

Response Variable

Out-of-hospital systems are frequently evaluated based on the time taken to respond to emergency requests and to transport patients to the hospital. The mean for the response variable is 29.67 min. and the median is 28.00 min.





Previous Research and Programs

- The performance of MLR and QR on patient characteristics and hospital transport times
- The performance of MLR and QR with the length of hospitalization for patients with gastrointestinal tract cancers
- We use Excel and R programming to fit these regression models to our data

Methods

Multiple Linear Regression (MLR):

MLR explains the relationship between a dependent variable and two or more predictors

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \varepsilon_i \quad (1)$$

Quantile Regression (QR):

QR focuses on approximating the conditional median or other quantiles of the response variable

$$Q_{y_i} = \beta_0 + \beta_i x_i + F_u^{-1}(\tau) \quad (2)$$

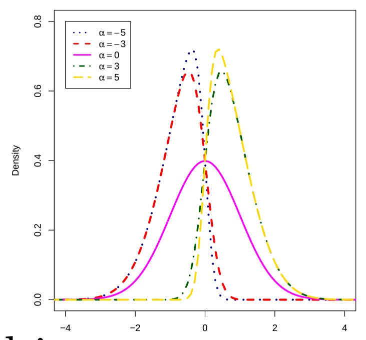
where F_u is the common distribution function.

Skew-Normal Regression (SNR):

SNR assumes that the random component in the regression model follows a continuous probability distribution known as skew-normal distribution, which takes into account skewness of the data.

$$f(y;\alpha) = 2\phi(y)\Phi(\alpha y); -\infty < y < \infty \quad (3)$$

where $\alpha \in \mathbb{R}$ is the shape parameter



The SNR model is

$$y_j = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \alpha z_j + \varepsilon_j$$

where $\alpha z_j + \epsilon_j \sim SN(0, \sigma^2, \alpha)$

Results on Univariate Models

When the transport interval was regressed on the type of pain:

MLR: $\hat{y_1} = 36.0699 - 6.9463x_1 - 6.9480x_2$

QR ($\tau = .5$): $\hat{y_1} = 35.0000 - 7.0000x_1 - 9.0000x_2$

SNR: $\hat{y_1} = 34.7574 - 5.1483x_1 - 6.3373x_2$

Transport interval was significantly associated in all three cases ($p < 0.001$).

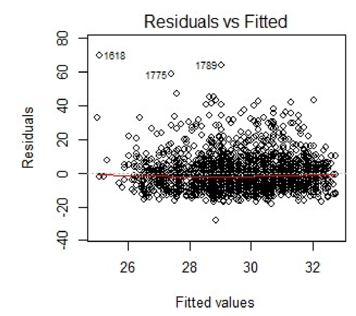
When the transport interval was regressed on age:

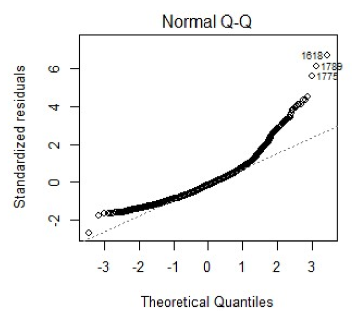
LR: $\hat{y_1} = 22.742 + 0.081x_1$

QR ($\tau = .5$): $\hat{y_1} = 22.442 + 0.093x_1$

SNR: $\hat{y_1} = 24.375 + 0.0925x_1$

Transport interval was significantly associated in all three cases ($p < 0.001$).





Results on Multivariate Models

MLR Multivariate Results:

Coefficient	Factor	Estimate	p-value
Chest Pain	Type	-5.1853	0.000
Trauma	Type	-5.4505	0.000
Fall Creek	Zone	17.452	0.000
Lincoln	Zone	21.483	0.000
Ludington	Zone	34.809	0.000
Pleasant Valley	Zone	16.174	0.000
Station 5	Zone	-11.615	0.000
Station 9	Zone	-6.9403	0.000
Sacred Heart	Dest.	-1.3352	0.000
Age		0.03772	0.000
M	Gender	-1.0896	0.0018
Early Morning	TOD	2.29633	0.0008
Late Night	TOD	2.06688	0.0006

QR 50th Quantile Multivariate Results:

Coefficient	Factor	Estimate	p-value
Chest Pain	Type	-4.5502	0.0022
Trauma	Type	-7.8421	0.000
Fall Creek	Zone	17.0430	0.000
Lincoln	Zone	27.3779	0.000
Ludington	Zone	40.1818	0.000
Pleasant Valley	Zone	16.2488	0.000
Station 5	Zone	-11.4497	0.000
Station 9	Zone	-7.4784	0.000
Mayo Clinic	Dest.	-5.37321	0.000
Sacred Heart	Dest.	-1.09569	0.000
St Josephs	Dest.	24.01914	0.000
Inter. Trans.	Urgency	28.36364	0.000
Age		0.04306	0.0036
M	Gender	-1.26794	0.010

SNR Multivariate Results:

Coefficient	Factor	Estimate	p-value
Chest Pain	Type	-3.706104	0.000
Trauma	Type	-5.331400	0.000
Drammen	Zone	24.596814	0.000
Fall Creek	Zone	17.023872	0.000
Lincoln	Zone	16.610279	0.000
Ludington	Zone	32.839908	0.000
Pleasant Valley	Zone	15.574052	0.000
Station 5	Zone	-10.144275	0.000
Station 9	Zone	-6.272788	0.000
Sacred Heart	Dest.	-1.171155	0.000
Age		0.057596	0.000
M	Gender	-0.928446	0.001
Early Morning	TOD	2.242161	0.000
Late Night	TOD	1.551187	0.002

Disagreements between the three regressions:

- 14 disagreements between MLR and QR
- 9 disagreements between MLR and SNR
- 9 disagreements between QR and SNR

AIC

Akaike Information Criterion (AIC):

An index to measure the fit of the model and compare the proposed models to other competitive models

$$AIC = 2k - 2ln(L)$$

The AIC of the three models are:

- AIC from MLR: 12457.44
- AIC from QR: 12192.63
- AIC from SNR: 12171.57

Discussions

Generally the univariate models provide better estimates in regards to regression analysis than the multivariate models. This is shown by the lower p-values that the univariate models produce, as well as in the lack of systematic pattern in the residual plots of the univariate models. Focusing on the univariate model comparing age and interval time, we observe that the effect of age on the median percentile, mean transport intervals, and patients with skewed transport intervals have estimates that are not adjusted for the other covariates.

According to the AIC values, the skew-normal regression gives a better fit to the dataset since it has a lower AIC value. Thus our hypothesis that is based off of the skewed response variable is correct. Quantile regression, linear regression, and skew-normal regression overall do not supply the same answers in regards to finding associations between out-of-hospital transport time and patient system characteristics because the regressions focus on responding to different questions. Some of the coefficient categories were statistically significant to the median quantile while others were significant to the mean transport times, and others to skewed transport times.

Limitations and Future Directions

Limitations to our research project include:

- Gridlock as a missing variable
- Alternative quantile regression models for analyzing out-of-hospital intervals
- Some category variables had many categories

Some future directions for this project include:

- Model assessments such as cross validation
- Analysis of percentiles other than the 50th

Acknowledgements

- Professor Mohammad Aziz for his support and guidance in this project
- The University of Wisconsin-Eau Claire for sponsoring this project
- The Office of Research and Sponsored Programs for the financial support
- UWEC Learning and Technology Services for printing this poster
- The Eau Claire Fire Department for the data

References

[1] Austin, P. C. & Schull, M. J. Quantile regression: A statistical tool for out-of-hospital research Acad. Emerg. Medicine 10, 789-797 (2003).

[2] Pourhoseingholi, M. A. et al. Comparing linear regression and quantile regression to analyze the associated factors of length of hospitalization in patients with gastrointestinal tract cancers Italian J. Public Heal. 6, 136-139 (2009)