

1100090952

Pusat Pembelajaran Digital Sultanah Nur Zahirah
Universiti Malaysia Terengganu.



tesis
RC 641.7 .T5 N8 2013



1100090952
Pemodelan dan perbandingan di antara model regresi poisson
dan model regresi poisson sifar-melambung terhadap penyakit
Thalassemia di kalangan kanak-kanak / Nur Syabiha Zafakali.

PUSAT PEMBELAJARAN DIGITAL SULTANAH NUR ZAHIRAH
UNIVERSITI MALAYSIA TERENGGANU (UMT)
21030 KUALA TERENGGANU

1100090952

1100090952		

Lihat Sebelah

HAK MILIK
PUSAT PEMBELAJARAN DIGITAL SULTANAH NUR ZAHIRAH

**PEMODELAN DAN PERBANDINGAN DI
ANTARA MODEL REGRESI POISSON DAN
MODEL REGRESI POISSON SIFAR-
MELAMBUNG TERHADAP PENYAKIT
THALASSEMIA DI KALANGAN KANAK-
KANAK**

NUR SYABIHA BT ZAFKALI

**IJAZAH SARJANA SAINS
UNIVERSITI MALAYSIA TERENGGANU**

2013

**PEMODELAN DAN PERBANDINGAN DI
ANTARA MODEL REGRESI POISSON DAN
MODEL REGRESI POISSON SIFAR-
MELAMBUNG TERHADAP PENYAKIT
THALASSEMIA DI KALANGAN KANAK-
KANAK**

NUR SYABIHA BT ZAFKALI

**Tesis Dihantar bagi Memenuhi Keperluan
Untuk Ijazah Sarjana Sains di Fakulti Sains dan
Teknologi Universiti Malaysia Terengganu**

2013

Abstrak tesis yang dikemukakan kepada Senat Universiti Malaysia Terengganu
sebagai memenuhi keperluan untuk ijazah Sarjana Sains

**PEMODELAN DAN PERBANDINGAN DI ANTARA MODEL REGRESI
POISSON DAN MODEL REGRESI POISSON SIFAR-MELAMBUNG
TERHADAP PENYAKIT THALASSEMIA DI KALANGAN KANAK-
KANAK**

NUR SYABIHA ZAFAKALI

2013

Penyelia Utama : Profesor Madya Wan Muhamad Amir W Ahmad, Ph.D
Penyelia Bersama : Nurfadhlina Abdul Halim, Ph.D
Fakulti : Sains dan Teknologi

Thalassemia merupakan penyakit gangguan darah yang diwarisi daripada keturunan. Jumlah pesakit yang menghidap thalassemia terutama di kalangan kanak-kanak telah dilaporkan meningkat dari tahun ke tahun dan telah dikenalpasti mendahului penyakit keturunan lain di kebanyakan bahagian dunia terutamanya di Malaysia. Selain daripada itu, kebanyakan pesakit turut terdedah kepada penyakit kronik lain seperti penyakit berdarah, mengalami masalah kesihatan, kegagalan jantung, influenza, anemia, pneumonia, bronkitis akut, asma, tonsillitis akut, jaundis dan tuberkulosis. Data pesakit yang terdiri daripada kalangan kanak-kanak telah dikumpulkan dan dijelmakan dalam bentuk statistik untuk dianalisis. Data tersebut dikategorikan mengikut skala-skala tertentu bersesuaian dengan analisis yang dijalankan. Disebabkan oleh kewujudan nilai sifar yang banyak dalam data, maka data ini dikatakan mengalami satu keadaan yang disebut sukatan serakan berlebihan. Permasalahan ini boleh diatasi melalui penggunaan kaedah regresi Poisson Sifar-Melambung. Kaedah penyelidikan yang terlibat dalam keadaan ini adalah model regresi Poisson dan model regresi Poisson Sifar-Melambung. Kaedah ini digunakan secara meluas untuk menganalisis data yang jenisnya berbentuk bilangan. Taburan Poisson diaplikasikan apabila nilai min dan nilai varians adalah sama. Manakala

model regresi Poisson Sifar-Melambung diaplikasikan apabila nilai min lebih kecil daripada nilai varians. Dapatan kajian menunjukkan bahawa, peratus pesakit daripada kalangan lelaki lebih tinggi berbanding perempuan dan pesakit ini terdiri daripada bangsa Melayu (91.0%), Cina (6.1%), India (0.4%) dan lain-lain kaum (2.4%). Daripada hasil penelitian, didapati bahawa pesakit thalassemia mempunyai hubungan yang signifikan terhadap pemboleh ubah peramal seperti masalah kesihatan 1.0768 ($p = 0.0001$), kegagalan jantung 0.9911 ($p = 0.0001$), influenza 0.5576 ($p = 0.0012$), anemia 0.3868 ($p = 0.0001$), pneumonia 0.6962 ($p = 0.0001$), bronkitis akut 0.7090 ($p = 0.0001$), asma 0.8172 ($p = 0.0001$), tonsillitis akut 0.5715 ($p = 0.0002$), jaundis 1.7287 ($p = 0.0001$) dan tuberkulosis 1.0139 ($p = 0.0002$). Namun begitu, bagi pemboleh ubah penyakit berdarah pula didapati bahawa nilai p adalah lebih besar daripada aras keertian dan secara kesimpulannya, pemboleh ubah ini tidak signifikan terhadap kajian kes yang dijalankan. Berdasarkan pemilihan model terbaik, nilai p bagi ujian Vuong adalah positif dan signifikan secara statistik ($p < 0.0001$, $n = 930$). Oleh yang demikian, keputusan menunjukkan bahawa model regresi Poisson Sifar-Melambung lebih sesuai digunakan untuk menganalisis data pesakit thalassemia di kalangan kanak-kanak. Hal ini kerana, model regresi Poisson lebih sesuai digunakan kepada data yang mempunyai sukatan serakan sama manakala model regresi Poisson Sifar-Melambung digunakan kepada data yang mempunyai sukatan serakan berlebihan.

Abstract of thesis presented to the Senate of Universiti Malaysia Terengganu in fulfillment of the requirement for the degree of Master of Science

MODELLING AND COMPARING BETWEEN POISSON REGRESSION MODEL AND ZERO-INFLATED POISSON REGRESSION (ZIPR) MODEL OF THALASSEMIA DISEASE AMONG CHILDREN

NUR SYABIHA ZAFAKALI

2013

Main Supervisor : Profesor Madya Wan Muhamad Amir W Ahmad, Ph.D
Co-Supervisor : Nurfadhlina Abdul Halim, Ph.D
Faculty : Science and Technology

Thalassemia is an blood related illness though descendant. Increasing number of patients suffering from thalassemia, especially among children have been reported year by year and have been identified ahead of other hereditary diseases in many parts of the world especially in Malaysia. Therefore, the increasing numbers of this illness annually, attracts the interest among the researchers to put an extra effort in order to overcome this illness. Other than that, most patients were also exposed other chronic diseases such as hemorrhagic illness, health problems, heart failure, influenza, anemia, pneumonia, acute bronchitis, asthma, acute tonsillitis, jaundice and tuberculosis. Data from patients especially among children has been successfully collected and it has been manifested in the form of statistic for analysis. The collected data then has been categorized according to certain scales appropriate to the analysis conducted. Due to the existence of many zero values in the data, the data is said to be suffering from overdispersion. This problem can be overcome by means of a zero-inflated Poisson regression model to reduce the value of zero. These models involved in this research are the Poisson regression model and the zero-inflated Poisson regression model. This methods are widely used to analyze count data. Actually these models are part of class of models in generalized linear models (GLM). One requirement of the Poisson distribution model is that the mean is equal

to the variance. While the zero-inflated Poisson regression model applied when the mean is smaller than the variance. The results showed that, the percentage of male patients are slightly higher than the female patients and patients consists of Malays (91.0%), Chinese (6.1%), Indians (0.4%) and other races (2.4%). Apart from the research done, it showed that the thalassemia patients have a significant relation to the predictor variables such as health problems 1.0768 ($p = 0.0001$), heart failure 0.9911 ($p = 0.0001$), influenza 0.5576 ($p = 0.0012$), anemia 0.3868 ($p = 0.0001$), pneumonia 0.6962 ($p = 0.0001$), acute bronchitis 0.7090 ($p = 0.0001$), asthma 0.8172 ($p = 0.0001$), acute tonsillitis 0.5715 ($p = 0.0002$), jaundice 1.7287 ($p = 0.0001$) and tuberculosis 1.0139 ($p = 0.0002$). However, for variable of hemorrhagic illness was found that the value of p is greater than the significant level and in conclusion, this variable is not significant to the case studies conducted. Based on the selection of the best model, the value of p for Vuong test is positive and statistically significant ($p < 0.0001$, $n = 930$). Thus, the final result from the analysis showed the zero-inflated Poisson regression model is more suitable in order to analyse the data of thalassemia patients among children. This is because, Poisson regression model is most appropriate for data that has no overdispersion while zero-inflated Poisson regression model is most appropriate for data that has overdispersion. This research can be used as the strategy of a better health management especially in patient management, to cut down the number of thalassemia patients among children.