RINGKASAN

Model regresi yang diperoleh dari metode kuadrat terkecil biasa (*Ordinary Least Square*/OLS) merupakan model regresi yang menghasilkan penaksir linier yang tidak bias yang terbaik (*Best Linear Unbiased Estimator*/BLUE) jika dipenuhi asumsi klasik. Asumsi klasik tersebut adalah nonmultikolinieritas, homoskedastisitas, nonotokorelasi, nilai rata-rata galat (*error*) populasi pada model stokastik sama dengan nol, variabel independen adalah nonstokastik (nilai konstan pada setiap kali percobaan yang dilakukan secara berulang) dan galat (*error*) berdistribusi normal.

Data dalam bidang kesehatan sering tidak memenuhi asumsi klasik. Masalah tersebut dapat diatasi dengan metode regresi yang tidak lagi terikat pada berbagai asumsi seperti metode pohon regresi dan metode regresi Cox.

Penelitian ini merupakan kajian statistik dalam membandingkan hasil analisis data survival dengan menggunakan metode regresi pohon dan metode regresi Cox. Penelitian ini diaplikasikan pada data lama perawatan penyakit diare pada balita di Rumah Sakit Dr. Soetomo Surabaya dengan variabel dependen adalah lama rawat dan variabel independen yaitu umur, jenis kelamin, dehidrasi, penyakit penyerta, lama diare, muntah dan status gizi.

Banyak data dalam penelitian ini adalah sebanyak 586 data. 57 data merupakan data yang tersensor dan 529 data merupakan data yang tidak tersensor.

Berdasarkan hasil analisis metode pohon regresi, variabel independen yang mempengaruhi lama perawatan penyakit diare pada balita di rumah sakit adalah lama diare, penyakit penyerta, status gizi dan dehidrasi. Karakteristik lama perawatan penyakit diare pada balita di rumah sakit dapat dikelompokkan ke dalam 7 kelompok yaitu :

1. Balita dengan lama diare sebelum dibawa ke rumah sakit lebih dari 13,5 hari maka rata-rata lama perawatannya di rumah sakit selama 10,556 hari.
2. Balita dengan lama diare sebelum dibawa ke rumah sakit kurang atau sama dengan 1,5 hari dan menderita penyakit Bronkopnemoni, Ensefalitis, Ensefalopati, Anemia, Marasmus dan penyakit lain-lain (bedah) maka rata-rata lama perawatannya 5,5 hari.
3. Balita yang menderita penyakit Bronkopnemoni, Ensefalitis, Ensefalopati, Anemia, Marasmus dan penyakit lain-lain (bedah) dengan $1,5 <$ lama diare $\leq 13,5$ hari maka lama perawatan balita di rumah sakit adalah 9,2 hari.
4. Balita dengan lama diare balita sebelum dibawa ke rumah sakit kurang atau sama dengan 13,5 hari dan menderita penyakit Meningitis, Malnutrisi, Kwarshiorkor, Marasmikwarshiorkor, Kejang Demam, Rhinopharyngitis dan tidak mempunyai penyakit penyerta, status gizi lebih dari 58,45 dengan dehidrasi sedang, berat dan tidak ada dehidrasi maka rata-rata lama perwatan balita di rumah sakit selama 5,857 hari
5. Lama diare balita sebelum dibawa ke rumah sakit kurang atau sama dengan 13,5 hari dan menderita penyakit Meningitis, Malnutrisi, Kwarshiorkor, Marasmikwarshiorkor, Kejang Demam, Rhinopharyngitis dan balita yang tidak mempunyai penyakit penyerta, gizi lebih dari 58,45 dan mengalami dehidrasi ringan maka rata-rata lama perawatan balita di rumah sakit selama 11 hari.

6. Balita yang menderita penyakit Meningitis, Malnutrisi, Kwarshiorkor, Marasmikwarshiorkor, Kejang Demam, Rhinopharyngitis dan balita yang tidak mempunyai penyakit penyerta dengan status gizi lebih dari 58,45 dengan lama diare kurang atau sama dengan 6,5 hari maka rata-rata lama perawatan balita di rumah sakit adalah 3,444 hari.
7. Balita yang menderita penyakit Meningitis, Malnutrisi, Kwarshiorkor, Marasmikwarshiorkor, Kejang Demam, Rhinopharyngitis dan balita tersebut tidak mempunyai penyakit penyerta dengan status gizi lebih dari 58,45 dengan $6,5 <$ lama diare $\leq 13,5$ hari maka rata-rata lama perawatan balita di rumah sakit adalah 4,625 hari.

Jika diperhatikan hasil metode pohon regresi pada kelompok 4 dan 5 , dimana untuk balita penderita diare dan mengalami dehidrasi ringan, lama rawatnya lebih lama dibandingkan dengan balita penderita diare yang mengalami dehidrasi sedang, berat atau tidak ada dehidrasi. Hal ini diduga akibat dari penyakit penyerta Marasmikwarshiorkor dan Rhinopharyngitis yang diderita oleh balita penderita diare.

Berdasarkan hasil metode regresi Cox, variabel independen yang mempengaruhi lama perawatan penyakit diare pada balita di rumah sakit adalah lama diare, penyakit penyerta dan status gizi. Untuk menduga lama perawatan diare balita ke-i di Rumah Sakit Dr. Soetomo digunakan model di bawah ini:

$$\widehat{h_i(t)} = \exp(-0{,}726\ \text{Bronkopnemoni}_{(i)} - 1{,}531\ \text{Ensefalitis}_{(i)} - 1{,}064\ \text{Ensefalopati}_{(i)} - 0{,}641\ \text{Penyakit Lain-lain (bedah)}_{(i)} - 0{,}086\ \text{Lama diare}_{(i)} + 0{,}008\ \text{Status gizi}_{(i)})\ \widehat{h_0(t)}$$

Berdasarkan model di atas, balita penderita diare yang menderita penyakit Bronkopnemoni, Ensefalitis, Ensefalopati atau penyakit lain-lain (bedah) mempunyai lama perawatan yang lebih lama di rumah sakit. Begitu juga balita penderita diare dimana balita tersebut mempunyai waktu yang lama sebelum dibawa ke rumah sakit maka lama perawatan balita tersebut lebih lama dibandingkan dengan balita penderita diare yang langsung dibawa ke rumah sakit. Semakin tinggi status gizi balita penderita diare semakin sedikit waktu yang diperlukan untuk perawatan balita tersebut di rumah sakit dan taksiran median waktu survival diperoleh pada balita penderita diare dengan lama perawatannya di rumah sakit selama 4 hari.

Kelebihan metode pohon regresi adalah metode ini dapat menduga langsung berapa lama perawatan penyakit diare pada balita di rumah sakit dan kelemahan dari metode pohon regresi adalah metode ini tidak dapat menentukan kategori mana dari data yang berskala nominal atau ordinal sangat berpengaruh di dalam model Kelebihan dari metode regresi Cox adalah dapat diketahui secara langsung tingkat signifikansi variabel yang berpengaruh di dalam model dan kelemahannya adalah metode ini tidak dapat secara langsung menentukan lama perawatan penyakit diare balita di rumah sakit.

Kesimpulan dari penelitian ini adalah variabel yang berpengaruh terhadap lama perawatan penyakit diare pada balita di Rumah Sakit Dr. Soetomo dengan menggunakan metode pohon regresi adalah variabel lama diare, penyakit penyerta, status gizi dan dehidrasi, sedangkan variabel yang berpengaruh terhadap lama

perawatan diare balita di rumah sakit dengan menggunakan metode regresi Cox adalah variabel penyakit penyerta, lama diare dan status gizi. Berdasarkan uraian di atas maka disarankan, jika ingin menduga secara langsung lama perawatan penyakit diare pada balita di rumah sakit dengan variabel berskala nominal atau ordinal berkategori sama dengan 2 (dua) maka gunakankanlah metode pohon regresi, tetapi jika ingin menduga lama perawatan penyakit diare pada balita di rumah sakit dengan menggunakan probabilitas dimana variabel berskala nominal atau ordinal berkategori lebih dari 2 (dua) maka gunakankanlah metode regresi Cox.

SUMMARY

Regression model obtained from the Ordinary Least Square (OLS) is a model that result in the Best Linear Unbiased Estimator (BLUE) if the classical assumptions are fulfilled. The classical assumptions are nonmulticollinearity, homoskedasticity, nonautocorrelation, in stchastic model the mean of error in the population is zero, the independent variable is nonstochastic (constant value in each repeated trial) and the error has normal distribution.

Health data often do not fulfill the classical assumption. Such problem may be overcome by applying regression method that is not bound to any assumptions, such as regression trees and Cox regression methods.

This study was a statistical investigation comparing results of survival data analysis by means of regression trees method and Cox regression method. This study was applied to the data on the length of hospitalization in underfives with diarrhea in Dr Soetomo Hospital, Surabaya. The dependent variable was the length of hospitalization and the independent variables were age, sex, dehydration, accompanying diseases, length of diarrhea, vomiting and nutritional status.

The number of data in this study was 586, 57 of which were censored data, and the remaining 529 were uncensored.

Based on analysis of regression trees method, the independent variables affecting length of hospitalization for underfives with diarrhea were length of diarrhea, accompanying diseases, nutritional status, and dehydration. Characteristics of the length of hospitalization for those underfives might be clssified into 7 groups as follows:

1. Underfives with pre-hospitalized length of diarrhea of more than 13.5 days had mean of hospitalization of 10.556 days.
2. Underfives with pre-hospitalized length of diarrhea of less than or equal to 1.5 days and suffered from bronchopneumonia, encephalitis, encephalopati, anemia, marasmus, and other (surgical) diseases had mean of hospitalization of 5.5 days.
3. Underfives with pre-hospitalized length of diarrhea of more than 1.5 days and less than and equal to 13.5 days and suffered from bronchopneumonia, encephalitis, encephalopati, anemia, marasmus, and other (surgical) diseases had mean of hospitalization of 9.2 days
4. Underfives with pre-hospitalized length of diarrhea of less than or equal to 13.5 days and suffered from meningitis, manutrition, kwarshiorkor, marasmikwashiorkor, febrile convulsion, rhinopharyngitis, no accompying diseases, nutritional status of more than 58.54, and had severe, moderate and no dehydration, had mean of hospitalization of 5.857 days.
5. Underfives with pre-hospitalized length of diarrhea of less than or equal to 13.5 days and suffered from meningitis, manutrition, kwarshiorkor, marasmicwashiorkor. febrile convulsion, rhinopharyngitis, no accompying diseases, nutritional status of more than 58.54, and had mild dehydration, had mean of hospitalization of 11 days.

6. Underfives with pre-hospitalized length of diarrhea of less than or equal to 6.5 days and suffered from meningitis, manutrition, kwarshiorkor, marasmikwashiorkor, febrile convulsion, rhinopharyngitis, no accompying diseases, and nutritional status of more than 58.54, had mean of hospitalization of 3.444 days.
7. Underfives with pre-hospitalized length of diarrhea of more than or 6.5 and less than or equal to 13.5 days and suffered from meningitis, manutrition, kwarshiorkor, marasmicwashiorkor, febrile convulsion, rhinopharyngitis, no accompying diseases, and nutritional status of more than 58.54, and had severe, moderate and no dehydration, had mean of hospitalization of 4.625 days.

Result of regression trees method in groups 4 and 5 can be seen, in those groups underfives with diarrhea and mild dehydration had a longer of hospitalzation than those with moderate, severe, or no dehydration. This might be due to the accompanying diseases, i.e., marasmicwarshiorkor and rhinopharyngitis, suffered by those infants.

Based on the result of Cox regression method, the independent variables affecting the length of hospitalization of underfives with diarrhea were length of diarrhea, accompanying diseases, and nutritional status. To estimate length of hospitalzation of i^{th} underfives with diarrhea at Dr Soetomo Hospital, the following model was used :

$$\mathbf{h_i(t) = \exp(-0.726\ bronchopnemonia_{(i)} - 1.531\ enchephalitis_{(i)} - 1.064\ encephalopathy_{(i)} - 0{,}641\ other\ (surgical)\ diseases_{(i)} - 0.086\ length\ for\ diarrhea_{(i)} + 0.008\ nutritional\ status_{(i)})\ h_0(t).}$$

Using above model, it was found that underfives with diarrhea and bronchopnemonia, encephalitis, encephalopathy, or other (surgical) diseases had longer length of hospitalization. Underfives with diarrhea with longer pre hospitalization also had longer length of hospitalization compared to those directly taken to the hospital. The higher the nutritional status, the shorter the length of their hospitalization. Estimated median of survival was obtained from underfives with length of hospitalization of 4 days.

The advantage of regression trees method is that it can directly estimate the length of hospitalization of those underfives, but it cannot determine category in data with nominal or ordinal scale that had hight influence in the model. In contrast, Cox regression model can determine directly the significance of the affecting variables in the model, while it cannot directly determine the length of hospitalization for underfives with diarrhea.

In conclusion, variables affecting length of hospitalization for underfives with diarrhea in Dr Soetomo Hospital using regression trees method are length of diarrhea, accompanying diseases, nutritional status, and dehidration; while using Cox regression method, those variables were accompanying diseases, length of diarrhea, and nutritional status. It can be recommended that regression trees method is appropriate to make direct estimation on the length of hospitalization for underfives with diarrhea using variables with nominal or ordinal scale and with the category of 2

(two). On the other hand, Cox regression method is appropriate to make estimation on the length of hospitalization for underfives with diarrhea using probability, in which the variables have nominal or ordinal scale with the category of more than 2 (two).

ABSTRACT

Regression analysis is one statistical tests commonly used for multifactorial data analysis. Regression analysis is interesting to use theoretically as it is based on mathematics. In addition, this analysis can be employed to investigate and model relations between variables. Regression analysis can be applied in variousdisciplines, such as engineering, physics, economics, management, biological and social sciences.

Problem that most commonly found in health data is that they do not fulfill the classical assumptions. Such problem may be overcome by means of regression analysis that is not bound to any assumptions, such as regression trees method and Cox regression analysis.

This study was aimed to survival data analysis (analyze factors affecting the length of hospitalization of underwifes with diarrhea) using regression trees method and Cox regression method. Data used in this study were those on the length of hospitalization of underfives with diarrhea in Dr Soetomo Hospital. The number of data in this study was 586, 57 of which were censored data, and the remaining 529 uncensored.

The cross validated relative error was 0.922 ± 0.041, the resubstitution relative error was 0.696 and the complexity was 45.497, the independent variables affecting the length of hospitalization in underfives with diarrhea, as revealed by means of regression trees method, were length of diarrhea, accompanying disseases, nutritional status, and dehydration, while variables affecting the length of hospitalization in underfives with diarrhea, as indicated by means of Cox regression method of $\alpha = 0.05$, were length of diarrhea, accompanying diseases, and nutritional status.

Keywords : *survival data analysis, regression trees method, Cox regression method.*