

RINGKASAN

Hubungan antara suatu variabel respon (*response variable*) yang bersifat kategori dengan satu atau beberapa variabel prediktor (*explanatory variable*) yang melibatkan sampel yang besar dianalisis dengan metode regresi logistik asimptotik. Namun bagaimana kalau ukuran sampelnya kecil atau datanya jarang (*sparse*), apakah metode regresi logistik masih relevan digunakan untuk analisis hubungan tersebut.

Penelitian ini merupakan suatu kajian statistik dalam penggunaan alternatif regresi logistik asimptotik, yaitu untuk melihat bagaimana aplikasi Regresi Logistik Eksak dalam menangani sampel kecil atau data yang jarang (*sparse*).

Dalam menangani sampel kecil berukuran 10, 20 dan 29 dan *sparse* data dengan sampel berukuran 55. Metode regresi logistik eksak dapat memberikan nilai taksiran dan solusi dalam pengujian hipotesis bahwa variabel bebas (keterpaparan terhadap informasi, pendidikan, pekerjaan, jumlah anak hidup yang dimiliki responden, pengetahuan, sikap dan peran serta masyarakat) secara bersama-sama berpengaruh terhadap variabel respon dalam hal ini status imunisasi dengan model

atau bentuk hubungan terbaik untuk ukuran sampel 20 adalah $\text{Log} \frac{\pi_j}{1 - \pi_j} = 5,1333$

+ 2,3582 Sikap, untuk ukuran sampel 29 adalah $\text{Log} \frac{\pi_j}{1 - \pi_j} = 2,9047$ sikap dan

untuk ukuran sampel 55 adalah $\text{Log} \frac{\pi_j}{1 - \pi_j} = -3,3608$ Pengetahuan + 1,5010 Sikap.

Demikian juga dengan interpretasi terhadap koefisien pembentuk model dimana Odds Rasio pada besar sampel 20, 29 dan 55 menunjukkan peluang terjadinya status imunisasi lengkap untuk ibu anak baduta yang mempunyai sikap mendukung masing-masing 10,5724 kali, 18,2595 kali dan 4,4862 kali jika dibandingkan dengan ibu anak baduta yang mempunyai sikap yang tidak mendukung.

Pengujian atas parameter dari model regresi logistik eksak baik sampel kecil maupun *sparse* data memberikan penolakan hipotesis nihil lebih besar (nilai p lebih kecil) dan memberikan nilai p-mid. Sedangkan Regresi logistik asimptotik tidak dapat memberikan solusi dengan p-mid.

Metode regresi logistik Eksak dengan *Conditional scores*nya memberikan nilai taksiran yang lebih baik daripada Metode regresi logistik asimptotik dengan uji Rasio likelihood, Wald dan Scores, dengan interval yang lebih melebar, sehingga peluang interval mengandung parameter koefisien regresi lebih besar dan dalam menentukan kemungkinan nilai pengamatan dalam distribusi bersyarat eksak dibangkitkan dengan algoritma yaitu *Multivariate Shift Algorithm* atau MSA.

Disarankan agar menggunakan metode Regresi logistik Eksak bila ukuran sampelnya kecil dan data yang jarang (*sparse*). Selain itu dikembangkan pula pada penggunaan data yang tidak seimbang (*imbalance data*) dan kasus dengan jumlah variabel yang berbeda-beda.

ABSTRACT

Correlation between a categorical response variable and one or several predictor variables involving large samples is analyzed using logistic regression method. However, if the sample size is small or the data is sparse, the relevance of conventional (asymptotic) logistic regression method to use in such correlation analysis will be questioned.

This is a statistical study for application of exact logistic regression in the condition of small sample size and sparse data.

Exact logistic regression analysis was done to cases with sample size of 10, 20, 29, and 55 taken from results of random sampling data obtained by Immunization. The dependent variable was immunization status, and the independent variable was people's exposure to information, education, occupation, living children, knowledge, attitude and participation. The data were analyzed with Logxact Turbo program.

Results showed that parameter estimation and hypothesis test using exact test provided better solution compared to conventional (asymptotic) regression test such as likelihood ratio, Wald, and score tests. Exact test also provided correlation type or model with sample size of 20, 30, and 55 and wider confidence interval compared to asymptotic inference type. Probability between those intervals had, therefore, larger parameter of population. For tests involving numerous independent variables, exact logistic regression method also provided better solution compared to three conventional (asymptotic) logistic regression tests.

Exact logistic regression method should be used in statistical test with small sample size or sparse data and tests that use table with sparse value in each of its cells.

Keywords: *exact logistic regression, asymptotic*