

## ABSTRAK

Tuberkulosis merupakan salah satu penyakit menular yang disebabkan oleh *Mycobacterium Tuberculosis* dan pada umumnya menyerang organ pernapasan. Penyakit tuberkulosis di Indonesia menempati urutan kedua dengan kasus tuberkulosis tertinggi di dunia dan Kabupaten Banyumas merupakan salah satu kabupaten dengan kasus yang cukup tinggi di Jawa Tengah. Penelitian ini bertujuan untuk menganalisis faktor-faktor yang mempengaruhi jumlah kasus tuberkulosis di Kabupaten Banyumas dengan menggunakan analisis regresi *count data*. Analisis regresi *count data* adalah metode statistika untuk mencari hubungan antara variabel respon yang menyatakan banyak peristiwa (*count variable*) dengan satu atau lebih variabel prediktor. Model regresi *count data* yang paling sederhana adalah regresi Poisson. Pada regresi Poisson, data diasumsikan berasal dari distribusi Poisson sehingga memenuhi sifat equidispersi, yaitu nilai *mean* sama dengan variansi. Namun, pada pengaplikasiannya asumsi tersebut sering tidak terpenuhi, misalnya terdapat kasus overdispersi. Overdispersi terjadi apabila dalam regresi Poisson nilai variansi lebih besar daripada nilai *mean*. Dalam penelitian ini, untuk mengatasi kasus overdispersi dilakukan pendekatan menggunakan *Generalized Poisson Regression* (GPR) dan regresi binomial negatif. Hasil penelitian menunjukkan bahwa data jumlah kasus tuberkulosis di Kabupaten Banyumas tahun 2019 mengalami overdispersi. Pemodelan data jumlah kasus tuberkulosis di Kabupaten Banyumas dengan model regresi binomial negatif lebih baik dibandingkan model GPR. Sementara itu, satu-satunya variabel prediktor yang berpengaruh terhadap jumlah kasus tuberkulosis di Kabupaten Banyumas yaitu rasio jenis kelamin usia produktif (15-49 tahun).

**Kata kunci:** tuberkulosis, regresi *count data*, overdispersi, *Generalized Poisson Regression* (GPR), regresi binomial negatif.

## **ABSTRACT**

*Tuberculosis is an infectious disease caused by Mycobacterium Tuberculosis and generally attacks the respiratory organs. Tuberculosis in Indonesia ranks second with the highest tuberculosis cases in the world and Banyumas Regency is one of the districts with quite high cases in Central Java. This study aims to analyze the factors that affect the number of tuberculosis cases in Banyumas Regency using regression analysis of count data. Regression analysis of count data is a statistical method to find the relationship between the response variable stating many events (count variable) with one or more predictor variables. The simplest count data regression model is Poisson regression. In Poisson regression, the data is assumed to be derived from the Poisson distribution so that meets the equidispersion property, that is, the mean value equal to the variance. However, in its application that assumption is often not fulfilled, for example, there are cases of overdispersion. Overdispersion occurs when in Poisson regression the variance value is greater than the mean value. In this study, to overcome the case of overdispersion, an approach was used using Generalized Poisson Regression (GPR) and negative binomial regression. The results showed that the data on the number of tuberculosis cases in Banyumas Regency in 2019 was overdispersion. The data modeling of the number of tuberculosis cases in Banyumas Regency with the negative binomial regression model is better than the GPR model. Meanwhile, the only predictor variable that affects the number of tuberculosis cases in Banyumas Regency is the sex ratio of productive age (15-49 years).*

**Keywords:** *tuberculosis, count data regression, Generalized Poisson Regression (GPR), negative binomial regression*