

# BAB I PENDAHULUAN

## 1.1 Latar Belakang Masalah

Analisis regresi merupakan ilmu statistika yang digunakan untuk mengetahui dan mempelajari hubungan peubah respon Y dengan satu atau lebih peubah prediktor X (Draper dan Smith, 1992). Regresi linier dibedakan menjadi dua bentuk, yaitu regresi linier sederhana dan regresi linier berganda. Dalam regresi linier sederhana hanya digunakan satu peubah prediktor, sedangkan pada regresi linier berganda menggunakan lebih dari satu peubah prediktor. Terdapat berbagai bentuk peubah respon pada model regresi, diantaranya peubah respon bersifat kuantitatif (rasio dan interval), dan peubah respon bersifat kualitatif atau kategorial.

Regresi logistik merupakan sebuah metode analisis statistik untuk menganalisa hubungan antara peubah prediktor dengan peubah respon yang mempunyai dua atau lebih kategori dengan peubah prediktor yang menggunakan skala kategorik maupun interval (Hosmer dan Lemeshow, 1989). Pendekatan regresi logistik digunakan karena metode tersebut dapat menjelaskan hubungan antara peubah respon dan peubah prediktor yang tidak dapat dijelaskan dengan regresi biasa.

Regresi logistik dengan peubah respon lebih dari dua kategori yang berskala nominal disebut juga regresi logistik multinomial (Fahrmeir dan Tutz, 1994), sedangkan jika peubah respon berskala ordinal digunakan regresi logistik ordinal. Melalui metode regresi logistik akan dihasilkan peluang untuk masing-masing kategori respon yang digunakan sebagai pedoman dalam pengklasifikasian.

Hampir sama dengan regresi logistik, regresi probit merupakan sebuah teknik analisis untuk mengestimasi kemungkinan sebuah peristiwa dengan peubah respon yang berskala nominal dan atau ordinal. Namun, berbeda dengan regresi logistik yang menggunakan fungsi logit, metode analisis regresi probit menggunakan fungsi distribusi normal (*Normal Distribution Function*). Kedua model hampir tidak dapat dibedakan penggunaannya. Hal ini didukung oleh Garson (1998) yang menyatakan bahwa model logistik dan probit menghasilkan nilai peluang hampir sama. Nilai peluang ini digambarkan dalam suatu kurva yang terdapat pada Lampiran 14.

Peneliti terdahulu yang telah mempelajari perbandingan model ini yaitu Saptadi (2006) yang membahas tentang model logistik dan probit dengan peubah respon biner dan Octavia (2007) dengan peubah respon politomus ordinal. Keduanya menyimpulkan bahwa model logistik dan probit sesuai digunakan pada peubah prediktor baik diskrit, kontinyu maupun campuran.

Garson (1998) menyatakan model logistik dan probit dapat digunakan pada peubah prediktor kategorial (diskrit) atau kontinyu. Namun, hingga saat ini belum ditemukan penelitian yang membandingkan kedua model dengan peubah respon bersifat multinomial. Oleh karena itu perlu dilakukan perbandingan antara model logistik dan probit dengan peubah respon multinomial, menggunakan peubah prediktor bersifat diskrit dengan ukuran sampel yang berbeda. Hal ini dimaksudkan agar diperoleh model terbaik dari keduanya.

Pendugaan parameter model regresi logistik dan probit dapat dilakukan dengan menggunakan metode *Maximum Likelihood Estimation* (MLE). Beberapa metode pemilihan model terbaik untuk fungsi MLE adalah *Goodness of fit Pearson*, *Deviance* dan  $R^2$  *Mc.Fadden*. Metode tersebut digunakan untuk data berukuran besar dengan jumlah kategori berbeda (Powell, 1997).

## **1.2 Perumusan Masalah**

Adapun perumusan masalah pada penelitian ini yaitu :

1. Bagaimana model regresi logistik dan regresi probit yang terbentuk pada data dengan peubah respon multinomial?
2. Manakah yang lebih baik antara regresi logistik dan regresi probit pada peubah respon multinomial?

## **1.3 Batasan Masalah**

Penelitian ini dibatasi pada model logistik dan probit peubah respon multinomial dengan lebih dari satu peubah prediktor dengan ukuran sampel lebih dari 50 yang berasal dari 10 data sekunder (penelitian). Metode yang dipakai untuk pendugaan parameter adalah metode kemungkinan maksimum (*Maximum Likelihood Estimation*) dengan iterasi *Newton-Raphson*. Uji *Pearson* digunakan untuk mendeteksi adanya nonmultikolinearitas.

#### **1.4 Tujuan Penelitian**

Tujuan penelitian ini adalah pembentukan model regresi logistik dan probit dengan peubah respon multinomial serta membandingkan keduanya berdasarkan nilai  $R^2$  *Mc.Fadden*.

#### **1.5 Manfaat Penelitian**

Manfaat yang dapat diperoleh dari hasil penelitian ini adalah mempelajari pembentukan model dan interpretasi, serta mengetahui model terbaik di antara model regresi logistik dan probit dengan peubah respon multinomial agar dapat digunakan dengan tepat dan maksimal dalam berbagai kasus yang ditemui.



UNIVERSITAS BRAWIJAYA

