

BAB I
PENDAHULUAN

1.1 Latar Belakang

Statistika adalah sekumpulan metode yang digunakan untuk merencanakan dan mengumpulkan data, menyajikan data, menganalisis data dan memberikan interpretasi, serta mengambil keputusan dan kesimpulan dalam situasi ketidakpastian dan variasi. Salah satu metode statistika yang sering digunakan dalam penyelesaian masalah pada kehidupan sehari-hari dengan variabel respon yang memiliki kategori adalah analisis regresi logistik. Analisis regresi logistik merupakan suatu analisis regresi untuk menjelaskan hubungan fungsional antara variabel respon (Y) yang bersifat dikotomis atau polikotomis dengan variabel-variabel prediktornya (X) (Hosmer dan Lemeshow, 2000). Pada analisis regresi logistik terdapat dua pendekatan untuk mengestimasi fungsi regresinya yaitu pendekatan parametrik dan pendekatan nonparametrik. Pendekatan regresi parametrik dapat digunakan apabila bentuk kurva fungsi regresinya telah diketahui misalnya berbentuk linier, kuadratik, kubik dan sebagainya (Abdy, 2019). Pada pendekatan regresi nonparametrik tidak mengasumsikan bentuk tertentu dari kurva fungsi regresinya. Fungsi regresi nonparametrik hanya diasumsikan *smooth* sehingga memberikan fleksibilitas yang tinggi dalam menentukan bentuk kurva fungsi regresinya (Eubank, 1999).

Pada pendekatan regresi nonparametrik, proses pengestimasi model regresi dengan menggunakan teknik *smoothing* tertentu, sehingga hasil estimasi model regresi yang diperoleh lebih fleksibel sesuai dengan pola data. Terdapat beberapa teknik *smoothing* dalam regresi nonparametrik, diantaranya estimator kernel, estimator lokal linier, estimator polinomial lokal, estimator *least square spline*, estimator *penalized spline*, dan estimator *deret fourier*. Diantara beberapa teknik *smoothing* tersebut, estimator *least square spline* adalah metode estimasi dengan model polinomial yang bersifat tersegmen atau terpotong-potong yang

mulus yang menggunakan basis fungsi *truncated* dan mengestimasi parameternya dengan meminimumkan jumlah kuadrat error. Kelebihan penggunaan estimator *least square spline* adalah memiliki fleksibilitas tinggi sehingga dapat mengatasi perubahan pola perilaku data yang menunjukkan naik atau turun yang tajam dengan bantuan titik-titik knot, serta kurva yang dihasilkan relatif mulus (Utami, 2018).

Pada kehidupan sehari-hari terdapat beberapa kasus yang melibatkan model regresi logistik dengan variabel respon yang terdiri dari dua kategorik (biner) atau lebih dari dua kategorik yang memiliki tingkatan (ordinal). Penelitian sebelumnya terkait penggunaan metode regresi logistik dengan variabel respon dikategorikan biner telah dilakukan oleh beberapa peneliti, di antaranya adalah penelitian Suliyanto *et al* (2020) mengenai penyakit diabetes dengan estimator lokal linier menghasilkan ketepatan klasifikasi pendekatan regresi logistik parametrik sebesar 80,2% dan pendekatan regresi logistik nonparametrik sebesar 100%. Penelitian lain dengan menggunakan estimator *spline* dilakukan oleh Rizka dan Chamidah (2019) mengenai tumor pada manusia menghasilkan ketepatan klasifikasi dengan pendekatan parametrik sebesar 80% sedangkan dengan pendekatan nonparametrik sebesar 85% dan penelitian Amalia *et al* (2020) mengenai resiko penyakit hipertensi menghasilkan ketepatan klasifikasi dengan pendekatan parametrik sebesar 65,6% dan pendekatan nonparametrik sebesar 90,7%. Dari hasil penelitian yang telah disebutkan dapat disimpulkan bahwa dengan menggunakan pendekatan regresi logistik nonparametrik, nilai ketepatan klasifikasinya lebih baik dibandingkan dengan pendekatan regresi logistik parametrik serta berdasarkan penelitian tersebut yang telah mengklasifikasikan penyakit dengan regresi logistik nonparametrik didapatkan estimator *spline* menghasilkan ketepatan klasifikasi tinggi dengan akurasi lebih dari 80 %.

Model regresi logistik dengan variabel respon lebih dari dua kategori dan dalam pengkategorianya memiliki tingkatan atau dapat diurutkan adalah regresi logistik ordinal. Penelitian sebelumnya yang menggunakan regresi logistik ordinal dengan pendekatan nonparametrik adalah penelitian Hastie and Tibshirani (1987) menggunakan pendekatan *Generalized Additive Models* (GAM) dengan algoritma

local scoring. Namun penelitian tersebut masih membahas model logit regresi logistik ordinal nonparametrik secara umum belum menggunakan pendekatan teknik smoothing tertentu untuk mengestimasi fungsi regresinya. Penelitian ini akan dijabarkan teori terkait estimasi model regresi logistik ordinal nonparametrik berdasarkan estimator *least square spline* dan diterapkan pada data riil dengan membuat algoritma dan program pada OSS-R.

Hasil penjabaran teori mengenai model regresi logistik ordinal nonparametrik berdasarkan estimator *least square spline* akan diimplementasikan pada data permasalahan status gizi balita menurut berat badan per tinggi badan (BB/TB) di Kabupaten Probolinggo. Kabupaten Probolinggo salah satu kabupaten di Provinsi Jawa Timur yang memiliki presentase prevelensi kasus gizi buruk dan gizi kurang cukup tinggi sebesar 20,59% dan angka tersebut lebih tinggi dibandingkan angka gizi kurang dan gizi buruk pada tingkat nasional maupun provinsi (Kementerian Kesehatan RI, 2018). Permasalahan gizi kurang pada balita akan membawa dampak negatif terhadap pertumbuhan fisik maupun mental yang selanjutnya akan menghambat perkembangan kecerdasan anak (Rahim, 2014). Data permasalahan status gizi menurut BB/TB yang akan digunakan dalam penelitian ini dibagi menjadi tiga kategori yaitu gizi kurang, gizi normal dan gizi lebih dengan memasukkan faktor-faktor yang dianggap mempengaruhi yaitu berat badan dan panjang badan lahir, umur ibu, tinggi badan ibu, jumlah anggota keluarga dan pengetahuan ibu tentang gizi. Apabila bentuk modelnya diketahui diharapkan kasus gizi kurang pada balita dapat diantisipasi sejak awal dan dapat dilakukan langkah – langkah pencegahan.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan pada sub bab sebelumnya, diperoleh rumusan masalah sebagai berikut.

1. Bagaimana mengestimasi model regresi logistik ordinal dengan pendekatan nonparametrik multiprediktor berdasarkan estimator *least square spline*?

2. Bagaimana membuat algoritma dan program pada OSS-R untuk mengestimasi model regresi logistik ordinal dengan pendekatan nonparametrik multiprediktor berdasarkan estimator *least square spline*?
3. Bagaimana mengimplementasikan dan menganalisis model regresi logistik ordinal dengan pendekatan nonparametrik multiprediktor berdasarkan estimator *least square spline* pada data balita di Kabupaten Probolinggo?

1.3 Tujuan Penelitian

Tujuan penelitian ini adalah sebagai berikut.

1. Mengestimasi model regresi logistik ordinal dengan pendekatan nonparametrik multiprediktor berdasarkan estimator *least square spline*.
2. Membuat algoritma dan program pada OSS-R untuk mengestimasi model regresi logistik ordinal dengan pendekatan nonparametrik multiprediktor berdasarkan estimator *least square spline*.
3. Mengimplementasikan dan menganalisis model regresi logistik ordinal dengan pendekatan nonparametrik multiprediktor berdasarkan estimator *least square spline* pada data balita di Kabupaten Probolinggo.

1.4 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut.

1. Bagi Mahasiswa
Menambah wawasan dan dapat mengaplikasikan ilmu yang telah dipelajari terkait metode statistika khususnya pada metode regresi logistik ordinal nonparametrik estimator *least square spline*.
2. Bagi Masyarakat
Menambah wawasan mengenai faktor – faktor yang mempengaruhi status gizi pada balita sehingga hal tersebut menjadi perhatian lebih bagi orang tua dalam menjaga keseimbangan gizi pada balita.
3. Bagi Pemerintah

Sebagai pertimbangan dalam merancang strategi dan menentukan kebijakan terkait permasalahan status gizi menurut berat badan per tinggi badan khususnya di Kabupaten Probolinggo.

1.5 Batasan Masalah

Penulisan skripsi ini dibatasi oleh responden penelitian yaitu ibu yang memiliki anak usia balita dan melakukan penimbangan di puskesmas Glagah dan Paiton, Kabupaten Probolinggo.