

BAB I

PENDAHULUAN

Bab ini berisi latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, metode penelitian, dan sistematika penulisan.

1.1 Latar Belakang Masalah

Dalam statistika tidak terlepas dari elemen-elemen yang berhubungan dengan data, variabel, dan skala pengukuran. Pengukuran dapat didefinisikan sebagai proses sistematis untuk menilai dan membedakan sebuah objek yang diukur. Pengukuran diatur berdasarkan kaidah-kaidah yang telah ditentukan. Dengan berdasarkan kaidah yang berbeda dapat menaksir skala serta pengukuran yang berbeda pula. Terdapat empat tipe pengukuran atau skala yang digunakan dalam statistika, yaitu nominal, interval, ordinal dan rasio. Data nominal dan data ordinal dapat disebut sebagai data kategori [1].

Salah satu bahasan dalam statistika yaitu regresi yang diperkenalkan oleh seorang ilmuwan yang bernama Francis Galton. Analisis regresi merupakan studi yang menjelaskan tentang ketergantungan atau hubungan pada satu atau lebih variabel bebas (*independent*) terhadap variabel terikat (*dependent*). Dalam analisis regresi terbagi menjadi dua jenis regresi yaitu regresi linier yaitu regresi yang meramalkan variabel prediktor pada persamaan pangkat satu (linier), dan regresi nonlinier yaitu regresi yang meramalkan variabel prediktor pada persamaan pangkat dua (nonlinier). Ada banyak jenis regresi nonlinier yang telah digunakan, diantaranya regresi eksponensial, regresi kubik, regresi *invers*, dan regresi logistik. Regresi logistik didefinisikan sebagai regresi nonlinier yang digunakan untuk mencari hubungan antara variabel respon kategori (dikotomi) dan variabel penjelas (prediktor) yang kontinu atau kategori. Pada setiap model regresi pasti memiliki parameter yang harus ditaksir. Untuk menaksir sebuah parameter penulis memerlukan metode-metode untuk estimasi. Secara umum metode estimasi atau penaksiran merupakan pengukuran terhadap nilai parameter (populasi) dari data sampel yang diketahui [2].

Ada beberapa metode yang digunakan untuk mengestimasi estimator dari regresi nonlinier, yaitu metode *maximum likelihood*, metode momen, estimator Bayes, dan metode *Least Square*. *Maximum Likelihood Estimation* (MLE) merupakan metode untuk memperoleh estimator titik dengan cara memaksimalkan fungsi *likelihood*, metode ini cocok digunakan untuk memprediksi fungsi regresi logistik karena fungsi logistik tidak hanya memprediksi titik tetapi juga memprediksi peluang. Regresi logistik merupakan salah satu model nonlinier yang digunakan untuk mencari hubungan antara peubah respon kategori dengan salah satu atau lebih peubah penjelas yang kontinu atau kategori [3].

Seorang peneliti seringkali dihadapkan dengan data yang tidak memenuhi asumsi normalitas, heteroskedastisitas, autokorelasi, multikolinieritas, data yang minim, data yang menyimpang, dan data yang tidak memiliki asumsi pada distribusinya, maka untuk mengatasi hal itu bisa menggunakan salah satu metode *resampling* yaitu metode *bootstrap*. Metode *bootstrap* didasarkan pada ide sederhana sebagai berikut: Sekalipun kita tidak tahu F kita bisa memperkirakannya dari data dan menggunakan perkiraan \hat{F} sebagai pengganti dari F itu sendiri. Ide ini mengarah ke beberapa pendapat yang berkaitan dengan *bootstrap* yang menyimpang pada bagaimana tepatnya aproksimasi \hat{F} diperoleh. Dua area luas dari metode *bootstrap* adalah *bootstrap* parametrik dan *bootstrap* nonparametrik [4].

Telah banyak penelitian sebelumnya yang membahas atau menggunakan metode *bootstrap* diantaranya yaitu dari Cempaka Putri (2013) ini mengkaji tentang estimasi *mean square error* dengan metode *bootstrap* parametrik dan metode *bootstrap* nonparametrik pada sampel ukuran kecil [5], I G. A. Anom Yudistira (2015) membahas metode *bootstrap* parametrik dan nonparametrik digunakan sebagai metode alternatif untuk mendapatkan nilai dugaan indeks kemampuan proses dan galat bakunya, karena pada pendugaan atas indeks kemampuan proses (C_p) sering kali sulit untuk memenuhi asumsi kenormalan yang mutlak [6], dan A.O Bello dkk (2015) pada jurnal ini metode *bootstrap* digunakan untuk meresampling data kategori pada data HIV/AIDS [7]. Yang

kemudian data asli dan data sampel *bootstrap* ditaksir dengan menggunakan metode *maximum likelihood estimation*, kemudian hasil estimasinya dibandingkan. Maka dalam penulisan skripsi ini penulis akan membahas perbandingan estimasi parameter data kategori dengan metode *bootstrap* parametrik dan *bootstrap* nonparametrik. Sehingga penulis memberikan judul “**Metode *Bootstrap* Parametrik dan Nonparametrik Untuk Analisis Data Kategori**”.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, maka rumusan masalah yang akan menjadi pembahasan dalam penulisan skripsi ini, adalah sebagai berikut :

1. Bagaimana langkah-langkah metode *bootstrap* parametrik dan nonparametrik untuk analisis data kategori ?
2. Bagaimana penerapan metode *bootstrap* parametrik dan nonparametrik untuk analisis data kategori ?

1.3 Batasan Masalah

Adapun batasan masalah dari skripsi ini adalah sebagai berikut :

1. Model yang digunakan merupakan model regresi logistik.
2. Metode yang digunakan untuk estimasi parameter regresi logistik adalah metode *Maximum Likelihood Estimation (MLE)*.
3. Data yang digunakan dalam skripsi ini merupakan data sekunder yaitu data luas wilayah, jumlah penduduk, kepadatan penduduk, dan data kabupaten/kota di Provinsi Jawa Barat, dengan 4 variabel yaitu 3 variabel bebas X berupa data kategori dan 1 variabel respon Y merupakan variabel respon biner.

1.4 Tujuan

Berdasarkan latar belakang dan rumusan masalah yang telah dijelaskan sebelumnya, maka tujuan dari penelitian skripsi ini adalah sebagai berikut:

1. Menentukan bagaimana langkah-langkah metode *bootstrap* parametrik dan nonparametrik untuk analisis data kategori.
2. Untuk mengetahui bagaimana penerapan metode *bootstrap* parametrik dan nonparametrik untuk analisis data kategori.

Berdasarkan tujuan skripsi yang akan dicapai, maka skripsi ini diharapkan akan memberikan manfaat sebagai berikut :

1. Dapat menyelesaikan permasalahan untuk analisis data kualitatif dan kategori ketika sampel terbatas dalam bidang ekonomi, politik, dan lain-lain.
2. Dapat menyelesaikan permasalahan dengan bentuk data kategori, baik kategori nominal maupun ordinal, misalnya pada bidang medis, sosial dan lain-lain.

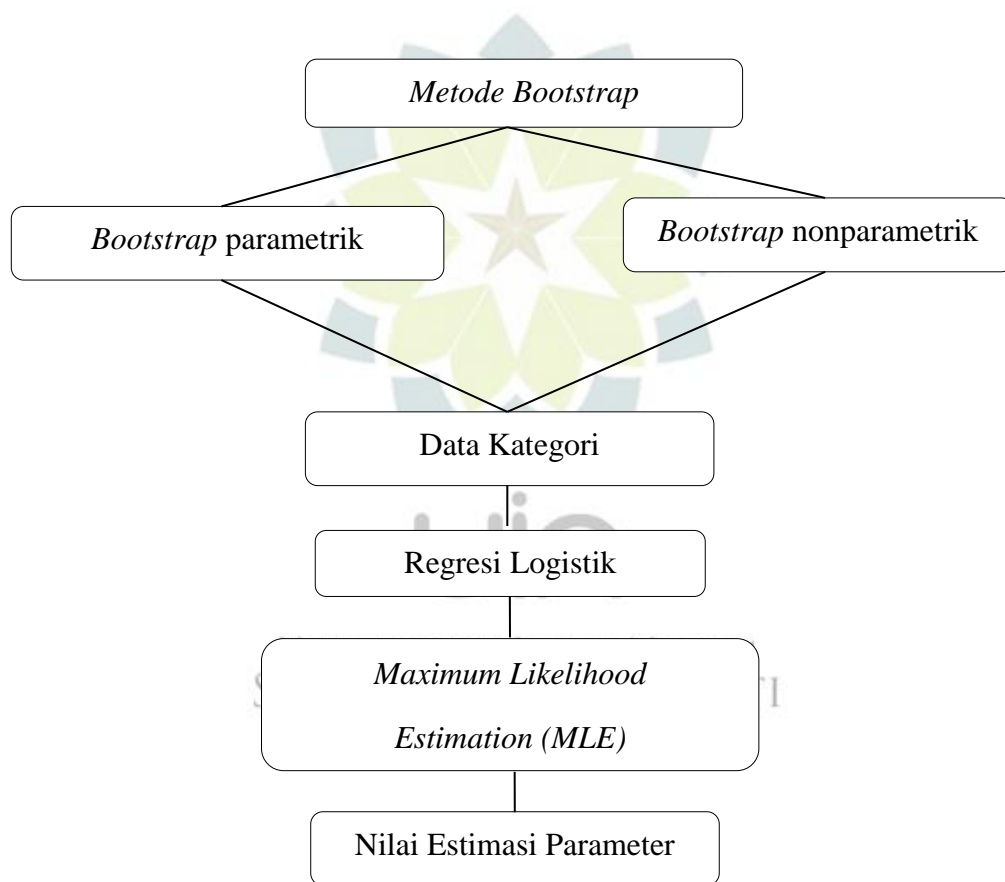
1.5 Metodologi Penelitian

Untuk tahap penelitian dalam skripsi ini dilakukan analisis dan penelitian dengan menggunakan metode *bootstrap* parametrik dan nonparametrik pada data kategori yang kemudian untuk setiap hasil *bootstrapping* dari masing-masing metode akan diestimasi dengan menggunakan metode *maximum likelihood estimation*. Selanjutnya hasil estimasi tersebut akan dibandingkan, sehingga penulis dapat mengetahui dengan metode mana hasil estimasi yang memiliki galat yang kecil.

Metode yang digunakan dalam penelitian ini adalah dengan mengkaji sumber pustaka berupa buku, jurnal, skripsi, dan tesis yang berkaitan dengan metode *bootstrap* parametrik dan metode *bootstrap* nonparametrik untuk analisis data kategori, dan pembahasan lain yang mendukung dalam penulisan skripsi ini. Data yang digunakan dalam penelitian ini adalah data sekunder yaitu data yang bukan hasil pengamatan sendiri atau data yang sudah ada, yaitu data luas wilayah, jumlah penduduk, kepadatan penduduk dan kabupaten/kota di Jawa Barat Tahun 2017.

Pengolahan data kategori yang berbentuk skala nominal dengan data yang sedikit. Model yang digunakan untuk data kategori ini adalah model regresi

logistik yang merupakan jenis regresi yang nonlinier. Untuk menaksir parameter regresi logistik penulis menggunakan metode *maximum likelihood estimation* (MLE), yaitu metode yang cocok digunakan untuk model regresi logistik dengan data kategori. Karena data yang digunakan dalam skripsi ini sedikit sehingga dilakukan replikasi untuk mengatasi data yang tidak memenuhi asumsi maka dalam skripsi ini menggunakan metode *bootstrap*, metode ini telah banyak digunakan dalam penelitian-penelitian terdahulu. Metode *bootstrap* ini memiliki dua pendekatan yaitu pendekatan secara parametrik dan nonparametrik. Ruang lingkup penelitian ini dapat dilihat pada Gambar 1.1 :



Gambar 1. 1 Kerangka Berpikir Metode *Bootstrap* Parametrik Dan Nonparametrik Untuk Analisis Data Kategori

1.6 Sistematika Penulisan

Untuk memahami lebih jelas tentang skripsi ini, pembahasan yang tertera pada skripsi ini dikelompokkan pada beberapa sub bab dengan sistematika penulisan sebagai berikut :

BAB I PENDAHULUAN

Bab ini berisi tentang latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, metode penelitian, dan sistematika penulisan.

BAB II LANDASAN TEORI

Bab ini berisi teori-teori yang melandasi pembahasan dalam skripsi ini. Secara garis besar, bab ini membahas tentang data kategori, model regresi logistik, metode *maximum likelihood estimation* dan metode *bootstrap* parametrik dan nonparametrik.

BAB III METODE *BOOTSTRAP* PARAMETRIK DAN NONPARAMETRIK UNTUK ANALISIS DATA KATEGORI

Bab ini berisi pembahasan utama dari skripsi ini, meliputi pembahasan mengenai langkah-langkah metode *bootstrap* parametrik dan nonparametrik dan bagaimana penerapan metode *bootstrap* parametrik dan nonparametrik untuk analisis data kategori.

BAB IV STUDI KASUS DAN SIMULASI PERHITUNGAN

Bab ini berisi penjelasan studi kasus sebagai contoh penerapan metode yang telah dijelaskan dalam Bab III. Analisis yang dilakukan mencakup interpretasi dari hasil penerapan tersebut dalam studi kasus.

BAB V PENUTUP

Bab ini berisi kesimpulan dari pembahasan Bab III dan analisis kesimpulan Bab IV serta saran penulis untuk pengembangan penelitian selanjutnya.