

BAB I

PENDAHULUAN

1.1 Latar Belakang

Statistika memiliki banyak kegunaan dalam kehidupan dan sangat luas dalam hal pengambilan keputusan yang tepat. Dalam kehidupan sehari-hari, sebenarnya telah dihadapkan dengan permasalahan sekaligus manfaat statistika. Contohnya adalah pada jajak pendapat yang disajikan oleh media-media dengan melakukan inferensi terhadap sampel yang ditarik. Berkaitan dengan sampel, statistika membutuhkan sampel dalam pengamatannya, baik sampel kecil maupun sampel besar, keduanya dapat memberikan hasil keputusan yang sesuai dengan masalah yang dikaji. Selain itu, tidak hanya dalam kehidupan sehari-hari, statistika juga sangat berguna dalam dunia pekerjaan. Salah satunya yaitu pada industri dengan menggunakan data sampel untuk memprediksi penjualan dan peluang mendapatkan keuntungan hasil produksi.

Begitu banyak kegunaan statistika yang telah direalisasikan dalam kehidupan nyata. Tidak asing, jika berbicara mengenai statistika maka identik dengan peluang atau probabilitas. Dalam statistika terdapat teori penting yang dijadikan dasar dan acuan, yaitu teori probabilitas. Objek utama pada teori probabilitas ialah peubah acak dan kejadian. Hubungan antara peubah acak dan kejadian menghasilkan probabilitas suatu kejadian. Sebagai dasar untuk statistika, teori probabilitas penting untuk kegiatan manusia yang melibatkan analisis kuantitatif data dalam jumlah yang besar. Pada peubah acak, terdapat peubah acak diskrit dan peubah acak kontinu yang memiliki karakteristik masing-masing. Peubah acak diskrit digunakan untuk menghitung data cacah dan peubah acak kontinu digunakan untuk menghitung data ukur. Seluruh kemungkinan nilai dan peluang dari suatu peubah acak dikenal sebagai distribusi peubah acak. Pada umumnya disebut dengan distribusi peluang atau distribusi probabilitas. Distribusi probabilitas secara garis besar dibagi menjadi dua, yaitu distribusi probabilitas diskrit dan distribusi probabilitas kontinu.

Salah satu distribusi peubah acak kontinu adalah Distribusi *Lindley* yang digunakan untuk memodelkan data waktu menunggu. Karakteristik Distribusi *Lindley* lebih fleksibel dibandingkan Distribusi Eksponensial, sehingga menunjukkan keunggulan dari distribusi tersebut [1]. Keadaan *overdispersi* sering kali menjadi kendala dalam memodelkan data, akan tetapi Distribusi *Lindley* tidak dapat digunakan untuk memodelkan data cacah/diskrit karena merupakan distribusi kontinu. Oleh karena itu, dilakukan diskritisasi pada Distribusi *Lindley* menggunakan metode yang mempertahankan fungsi *survival*. Distribusi hasil diskritisasi Distribusi *Lindley* biasa disebut Distribusi *Discrete Lindley* yang memiliki satu parameter dan dapat digunakan untuk memodelkan data *overdispersi* [2].

Data yang mengalami *overdispersi* merupakan salah satu kasus yang harus diselesaikan, dibutuhkan adanya penyesuaian dengan data, baik data cacah maupun data waktu hidup sehingga memberikan lebih banyak fleksibilitas untuk memodelkan data-data tersebut. Oleh karena itu, dilakukan langkah untuk mengeksponensialkan Distribusi *Discrete Lindley* yang menghasilkan Distribusi *Exponentiated Discrete Lindley*. M. El-Morshedy, M.S Eliwa dan H. Nagy memperkenalkan konsep *Exponentiated Discrete Lindley Distribution* yang merupakan perkembangan dari Distribusi *Discrete Lindley*, biasa disebut dengan EDLi. Distribusi EDLi memiliki dua parameter yang dapat disesuaikan dengan data. Hal itu memberikan lebih banyak fleksibilitas daripada Distribusi *Discrete Lindley* untuk memodelkan data yang mengalami *overdispersi*. Selain itu, Distribusi EDLi memberikan kesesuaian terbaik untuk data cacah dan data waktu hidup sehingga dapat digunakan untuk pemodelan data dalam analisis *survival*, reliabilitas dan waktu kegagalan [3].

Salah satu penelitian yang berkaitan dengan Distribusi *Lindley* adalah Distribusi *Lindley*-Eksponensial yang merupakan gabungan dari dua distribusi, yaitu Distribusi Eksponensial dan Distribusi *Lindley* [4]. Distribusi ini dapat digunakan untuk memodelkan data waktu menunggu, dimana data waktu menunggu merupakan salah satu contoh dari data antar kejadian, dan dapat menangani bentuk *hazard* yang berbeda-beda [4]. Adapun Distribusi Binomial

Negatif-*Lindley* yang mampu menangani masalah *overdispersi* dan menangani jika jumlah frekuensi yang banyak [5]. Selain itu, terdapat Distribusi *Log-Lindley* yang dapat digunakan untuk menangani permasalahan pada asuransi, khususnya mengenai distribusi risiko untuk menghitung nilai risiko asuransi yang diharapkan [6].

Begitu banyak penelitian pada bidang sains yang meliputi bidang Biologi, bidang Kesehatan maupun bidang Kedokteran yang menghasilkan data yang berhubungan dengan data waktu hidup dari suatu objek yang diteliti. Perlu diketahui bahwa data waktu hidup merupakan peubah acak non negatif. Salah satu kasus yang berkaitan dengan data waktu hidup dan sedang mengguncang kalangan masyarakat ialah COVID-19. COVID-19 adalah penyakit menular yang disebabkan oleh sindrom pernapasan akut Sars-CoV-2. Wabah penyakit ini begitu mengguncang masyarakat dunia, hingga hampir 200 negara terjangkit termasuk Indonesia.

Pada statistika, analisis yang digunakan untuk menganalisis data waktu hidup ialah analisis *survival*. Menurut Lawless (1943) data waktu hidup yang diperoleh dari suatu percobaan dapat berbentuk data lengkap apabila objek dalam percobaan diuji hingga semuanya “mati” [7]. Fungsi distribusi yang menganalisis data waktu hidup biasanya disebut sebagai fungsi distribusi tahan hidup yang didasarkan pada asumsi tertentu mengenai distribusi populasi. Berikut beberapa distribusi yang dapat digunakan untuk menganalisis data waktu hidup diantaranya Distribusi Ekspensial, Distribusi Gamma, Distribusi Weibull, Distribusi Rayleigh, dan lain-lain. Diantara beberapa distribusi tersebut, dalam skripsi ini dipilih Distribusi *Exponentiated Discrete Lindley* untuk mengetahui apakah distribusi dari data waktu hidup yang diasumsikan telah menggambarkan keadaan yang sesungguhnya. Langkah yang diperlukan untuk menganalisis fungsi distribusi terhadap data waktu hidup adalah dengan mengestimasi parameternya.

Setiap distribusi probabilitas memiliki parameter yang digunakan sebagai karakteristik pengukuran dari suatu distribusi. Begitu juga dengan Distribusi *Exponentiated Discrete Lindley* yang memiliki dua parameter. Pendugaan parameter adalah persoalan bagaimana cara menduga atau menguji hipotesis

parameter populasi yang belum diketahui dengan contoh acak dan perhitungan probabilitas. Harga parameter yang sebenarnya tidak diketahui akan ditaksir berdasarkan statistik sampel yang diambil dari populasi yang bersangkutan.

Pada umumnya, penaksir parameter dibagi menjadi dua bagian, penaksir titik dan penaksir interval. Beberapa metode yang digunakan untuk memperoleh penaksir titik, yaitu metode momen, Metode Kuadrat Terkecil (MKT), *Maximum Likelihood Estimation* (MLE) dan Metode Bayes. Penaksir parameter yang digunakan untuk menaksir parameter Distribusi *Exponentiated Discrete Lindley* adalah *Maximum Likelihood Estimation* (MLE).

Berdasarkan latar belakang diatas maka mendorong peneliti untuk mengadakan penelitian mengenai penerapan data waktu hidup pasien COVID-19 menggunakan Distribusi EDLi. Oleh karena itu, penulis memberikan judul pada skripsi ini adalah “Analisis Distribusi *Exponentiated Discrete Lindley* dengan Kendala Data *Overdispersi* pada Data Masa Hidup Pasien COVID-19”.

1.2 Rumusan Masalah

Rumusan masalah dalam skripsi ini adalah sebagai berikut:

1. Bagaimana membangun model Distribusi *Exponentiated Discrete Lindley*?
2. Bagaimana penerapan analisis Distribusi *Exponentiated Discrete Lindley* dengan kendala data *overdispersi* pada data masa hidup pasien COVID-19?

1.3 Batasan Masalah

Batasan masalah dalam skripsi ini adalah sebagai berikut:

1. Distribusi yang digunakan adalah Distribusi *Exponentiated Discrete Lindley* dengan dua parameter.
2. Metode estimasi parameter yang digunakan yaitu *Maximum Likelihood Estimation* (MLE) dengan pendekatan metode *Newton-Raphson*.
3. Data yang digunakan merupakan data lamanya perawatan 60 pasien COVID-19 di RSUD Kota Bandung.
4. Uji *overdispersi* yang digunakan adalah Uji *Deviance*.

5. Uji signifikansi parameter yang digunakan adalah Uji Rasio *Likelihood*.
6. Uji kecocokan distribusi yang digunakan adalah *Akaike Information Criterion*.

1.4 Tujuan Penelitian

Tujuan dari penulisan skripsi ini adalah sebagai berikut:

1. Menentukan model Distribusi *Exponentiated Discrete Lindley*.
2. Menerapkan analisis Distribusi *Exponentiated Discrete Lindley* dengan kendala data *overdispersi* pada data masa hidup pasien COVID-19.

1.5 Metode Penelitian

Langkah-langkah penyelesaian yang sesuai dengan tujuan penelitian adalah sebagai berikut:

1. Melakukan studi literatur terhadap berbagai referensi berkaitan dengan penelitian yang dilakukan bersumber dari buku, jurnal dan penelitian sebelumnya.
2. Membangun model Distribusi *Exponentiated Discrete Lindley* dengan cara *memproof* berdasarkan definisi dan teorema yang sudah tersedia pada referensi.
3. Menentukan estimasi parameter Distribusi *Exponentiated Discrete Lindley* pada data masa hidup pasien COVID-19 menggunakan *Maximum Likelihood Estimation* (MLE) dengan pendekatan metode *Newton-Raphson*.
4. Menguji *overdispersi* pada data masa hidup pasien COVID-19 menggunakan Uji *Deviance* dengan bantuan *R software*.
5. Menguji signifikansi parameter yang diperoleh menggunakan Uji Rasio *Likelihood* terhadap data masa hidup pasien COVID-19.
6. Menguji kecocokan model Distribusi *Exponentiated Discrete Lindley* menggunakan *Akaike Information Criterion* terhadap data masa hidup pasien COVID-19.
7. Menyusun algoritma berdasarkan langkah-langkah yang telah dibuat.

8. Membuat program komputer berdasarkan algoritma tersebut menggunakan *R Software*.
9. Menerapkan Distribusi *Exponentiated Discrete Lindley* dengan kendala data *overdispersi* pada data masa hidup pasien COVID-19, selanjutnya melakukan analisis dan hasil interpretasi.

1.6 Sistematika Penulisan

Berdasarkan sistematika penulisannya, skripsi ini terdiri atas lima bab, serta daftar pustaka dan lampiran, dimana setiap bab terdapat beberapa sub bab.

BAB I PENDAHULUAN

Bab ini berisi tentang pendahuluan diantaranya adalah latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, metode penelitian dan sistematika penulisan dari masalah yang dikaji.

BAB II LANDASAN TEORI

Bab ini berisi penjelasan tentang teori-teori yang berkaitan dengan masalah yang akan dikaji. Secara garis besar, bab ini mencakup semua yang berkaitan dengan Distribusi *Exponentiated Discrete Lindley* dengan kendala data *overdispersi*.

BAB III ANALISIS DISTRIBUSI *EXPONENTIATED DISCRETE LINDLEY* DENGAN KENDALA DATA *OVERDISPERSI*

Bab ini berisi pembahasan utama dari skripsi yang dikaji, meliputi pembahasan mengenai model Distribusi *Exponentiated Discrete Lindley* dan penerapan Distribusi *Exponentiated Discrete Lindley* pada data masa hidup pasien COVID-19.

BAB IV STUDI KASUS DAN ANALISIS

Bab ini berisi tentang studi kasus, simulasi perhitungan model menggunakan Distribusi *Exponentiated Discrete Lindley*, serta analisis dan hasil analisis.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dari seluruh pembahasan yang telah dikaji berdasarkan tujuan yang ingin dicapai. Selain itu, diberikan saran untuk pengembangan lebih lanjut terhadap topik pembahasan selanjutnya.

DAFTAR PUSTAKA

LAMPIRAN

