

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Statistika adalah cabang ilmu yang mempelajari teknik pengumpulan, penyajian, analisis, serta interpretasi data. Statistika memegang peranan penting sebagai disiplin ilmu terhadap berbagai bidang, terutama dalam sebuah penelitian. Statistik inferensial merupakan cabang ilmu yang mengkaji semua metode analisis terhadap beberapa data untuk kemudian membuat prediksi atau kesimpulan tentang sumber data tersebut [1]. Secara umum, ciri fenomena atau konsep tertentu yang menjadi fokus penelitian disebut populasi. Parameter merupakan indikator untuk mengukur populasi dengan menggunakan bermacam-macam ukuran, seperti rata-rata, median, varians, proporsi ataupun modus. Perlu diketahui bahwa tidak semua unsur dari sebuah populasi dapat diukur itu menyebabkan pengukuran populasi akan dirasa sulit, maka untuk memudahkan pengukuran akan digunakan sampel dari populasi tersebut. Setiap unsur dalam populasi yang membentuk sebuah sampel memiliki probabilitas tertentu. Hal ini menghasilkan elemen-elemen dari suatu populasi tersebar, sehingga membentuk distribusi tertentu. Penduga suatu parameter (estimator) haruslah bersifat tidak bias (nilai penaksir akan sama dengan atau mendekati nilai penduga), efisien (nilai variansi penduga kecil), serta konsisten (bertambahnya jumlah sampel tidak berbanding terbalik dengan nilai penaksir).[2]

Statistika memiliki banyak distribusi untuk menunjukkan model penyebaran data, salah satunya adalah Distribusi Pareto. Distribusi Pareto dikenalkan oleh Vilfredo Pareto sekitar tahun 1848-1923. Dimana Vilfredo Pareto ini merupakan seorang insinyur, sosiolog, ekonom sekaligus seorang filsuf kebangsaan Italia. Buku Ekonomi tahun 1897 merupakan salah satu karya atas gagasan yang ia miliki [3]. Distribusi Pareto  $f(x; \alpha, \beta)$  termasuk distribusi kontinu yang memiliki dua parameter yaitu parameter bentuk ( $\alpha$ ) dan parameter skala ( $\beta$ ). Distribusi Pareto menjadi salah satu distribusi yang banyak digunakan dalam beberapa kasus, seperti bidang ekonomi, bisnis, asuransi, bahkan dalam bidang kesehatan. Distribusi Pareto merupakan salah satu keluarga dari distribusi *power law*, yang mana dalam

distribusi ini nilai pangkat dari *probability density function (PDF)* mempengaruhi terhadap nilai standar deviasi atau rataannya. Distribusi Pareto diklasifikasikan sebagai distribusi *heavy tailed*, dimana distribusi ini mampu mengilustrasikan data berpeluang besar ketika nilai pengamatannya besar. Kelemahan Distribusi Pareto terdapat pada kemampuannya dalam menggambarkan fungsi kepadatan peluang, Distribusi Pareto hanya memiliki sebuah bentuk grafik yang berbentuk eksponensial. Oleh karena itu, Distribusi Pareto dianggap kurang fleksibel dalam menggambarkan suatu data.[4]

Peristiwa yang muncul pada kehidupan nyata menunjukkan adanya variasi data (variabilitas) yang sangat besar yang secara fisik membentuk pola multimodal[5]. Fenomena yang demikian bila didekati dengan distribusi *single univariat* maka kemungkinan besar akan bersifat bias pada analisisnya[6]. Hal ini membawa pemikiran membentuk model *mixture* sebagai model alternatif lain untuk menyelesaikan variabilitas data. Dimana model *mixture* ini merupakan gabungan atas beberapa sub-populasi berdistribusi *single univariat*[7]. Namun, perlu dipahami bahwa tidak semua data akan lebih baik jika dimodelkan dengan bentuk *mixture*, karena justru akan lebih baik jika dimodelkan dengan *single univariat*. Membentuk pendekatan distribusi yang tepat akan berdampak pada analisis yang akurat, karena bila kurang sesuai maka akan cenderung menjadi bias.

Penerapan model *mixture* dapat dilakukan dalam berbagai bidang, salah satu contohnya kasus yang akan diterapkan dalam penelitian ini, yaitu Tingkat Pengangguran Terbuka (TPT). Tingkat pengangguran yang cukup tinggi menyebabkan masalah pembangunan ekonomi di Indonesia. Badan Pusat Statistik (BPS) mengekspos data Tingkat Pengangguran Terbuka (TPT). Dalam skala nasional TPT tercatat sebesar 6,49% pada tahun 2021 [8]. Menurut Survey Angkatan Kerja Nasional (SAKERNAS) tahun 2021 menyatakan rata-rata TPT Pulau Jawa sebesar 7,25%. Hal ini menjadi ironi, mengingat Pulau Jawa merupakan pusat pemerintahan yang dihuni lebih dari setengah penduduk Indonesia sehingga dijadikan pusat perekonomian nasional. Menurut Survey Penduduk Antar Sensus (SUPAS) BPS tahun 2020, jumlah penduduk di Pulau Jawa sebesar 151.591,2 ribu jiwa itu setara dengan 56,1 % dari total penduduk Indonesia [9]. Jumlah penduduk yang besar memiliki dampak positif yaitu semakin meningkatnya jumlah tenaga

kerja bagi pembangunan ekonomi di Indonesia. Namun dilain sisi, pertumbuhan penduduk yang besar ini pula dapat menimbulkan berbagai masalah dan hambatan bagi upaya pembangunan yang dilakukan karena tingginya jumlah angkatan kerja yang tidak diimbangi dengan meningkatnya lowongan pekerjaan di lapangan.[10]

Latar belakang ekonomi Pulau Jawa yang berkembang pesat dan beragam seiring dengan pertumbuhan penduduk yang tinggi, memberikan tantangan dalam mengelola tingkat pengangguran terbuka. Faktor-faktor seperti pertumbuhan ekonomi, kebijakan ketenagakerjaan, dan dinamika pasar kerja dapat mempengaruhi tingkat pengangguran di pulau ini. Meskipun pandemi COVID-19 sudah dapat tertangani tetapi dampaknya terhadap pasar kerja dan tingkat pengangguran masih perlu diperhatikan.

Berdasarkan penjelasan di atas, penulis tertarik untuk mengkaji lebih lanjut mengenai model Distribusi Pareto Campuran pada data sekunder Tingkat Pengangguran Terbuka (TPT) di Pulau Jawa pada tahun 2021 yang didapat dari Badan Pusat statistik (BPS). Estimasi parameter dilakukan dengan menggunakan *Maximum Likelihood Estimation (MLE)*, lalu disempurnakan untuk iterasi numerik menggunakan *Newton-Raphson (NR)*. Oleh karena itu, peneliti mengangkat skripsi berjudul **“Analisis Model Distribusi Pareto Campuran pada Tingkat Pengangguran Terbuka (TPT) di Pulau Jawa Tahun 2021”**.

## 1.2 Rumusan Masalah

Rumusan masalah pada pembahasan masalah di atas ini dapat diuraikan sebagai berikut :

1. Bagaimana bentuk model Distribusi Pareto Campuran
2. Bagaimana langkah-langkah estimasi parameter Distribusi Pareto Campuran
3. Bagaimana penerapan model Distribusi Pareto Campuran dalam kasus Tingkat Pengangguran Terbuka (TPT) di Pulau Jawa pada tahun 2021

### 1.3 Batasan Masalah

Terdapat beberapa batasan dalam pembahasan masalah ini, diantaranya sebagai berikut :

1. Penulis hanya menggabungkan dua distribusi dengan dua bobot yang berbeda.
2. Estimasi parameter menggunakan metode *Maximum Likelihood Estimation* (MLE) dan iterasi numerik *Newton-Raphson*.
3. *Software* yang digunakan dalam skripsi ini yaitu *R*, *SPSS*, dan *Easyfit*.
4. Penelitian ini tidak membahas tentang uji kecocokan model distribusi campuran lebih lanjut.
5. Penulis hanya membahas Distribusi Pareto Campuran menggunakan Distribusi Pareto tipe (I). Penulis tidak membahas Distribusi Pareto Campuran menggunakan Distribusi Pareto tipe (II), (III), dan (IV).
6. Data yang digunakan adalah data Tingkat Pengangguran Terbuka (TPT) di Pulau Jawa pada bulan Agustus tahun 2021 yang diperoleh dari *website* Badan Pusat Statistika Pusat (BPS).

### 1.4 Tujuan Penelitian

Berdasarkan rumusan masalah di atas, maka tujuan yang ingin dicapai pada penelitian ini adalah

1. Menentukan bentuk model Distribusi Pareto Campuran.
2. Menentukan langkah-langkah estimasi parameter dari Distribusi Pareto Campuran
3. Menerapkan model Distribusi Pareto Campuran pada data Tingkat Pengangguran Terbuka (TPT) di Pulau Jawa tahun 2021.

### 1.5 Metode Penelitian

Metode yang digunakan pada penelitian ini adalah metode kuantitatif. Metode ini berdasarkan pada studi literatur, buku, jurnal, skripsi, dan sumber lainnya sebagai sarana pengumpulan berbagai data baik berupa teori ataupun data statistik yang berkaitan dengan Distribusi Pareto Campuran. Kemudian, data

diperoleh dengan cara pengambilan data sekunder yang berasal dari *database* Badan Pusat Statistika (BPS). Data yang digunakan adalah data Tingkat Pengangguran Terbuka (TPT) Menurut Kabupaten/Kota se-Jawa tahun 2021. Selanjutnya, dilakukan proses analisis, yaitu menentukan estimasi parameter dalam Distribusi Pareto Campuran dengan menggunakan *Maximum Likelihood Estimation (MLE)* dilanjut dengan iterasi numerik menggunakan *Newton-Raphson (NR)*.

## **1.6 Sistematika Penulisan**

Penulisan skripsi ini disusun dengan menggunakan sistematika penulisan yang terdiri dari lima bab, dan setiap bab terdiri dari beberapa subbab sebagai berikut :

### **BAB I PENDAHULUAN**

Bab ini berisi pembahasan mengenai latar belakang masalah, rumusan masalah, batasan masalah, tujuan dan manfaat penelitian, metode penelitian, dan sistematika penulisan dalam skripsi ini.

### **BAB II LANDASAN TEORI**

Bab ini berisi penjelasan mengenai teori-teori yang menjadi landasan literatur guna menyelesaikan masalah pada BAB III. Secara garis besar, mencakup materi yang berkaitan dengan Distribusi Pareto Campuran untuk memahami istilah-istilah yang tercantum dalam kajian ini.

### **BAB III ESTIMASI PARAMETER DISTRIBUSI PARETO CAMPURAN**

Bab ini bersisi tentang penjelasan dan langkah-langkah membentuk model Distribusi Pareto Campuran, dan estimasi parameter Distribusi Pareto Campuran yang digunakan pada penelitian.

#### **BAB IV STUDI KASUS DAN ANALISA**

Bab ini berisi penerapan metode yang dijelaskan pada BAB III dengan menganalisis studi kasus sebagai contoh penerapan, yang dilakukan hingga menghasilkan interpretasi dari hasil penerapan.

#### **BAB V KESIMPULAN DAN SARAN**

Bab ini menjelaskan kesimpulan dari hasil pembahasan dan studi kasus penelitian yang telah dilakukan beserta saran untuk pengembangan lebih lanjut yang dapat dilakukan dari topik pembahasan tersebut.

