

BAB I

PENDAHULUAN

Pengetahuan tentang fungsi densitas sangat membantu dalam banyak hal. Salah satu contoh adalah untuk mengetahui bagaimana terjadinya suatu observasi dari sekumpulan observasi X , misalnya dalam interval $X=(a,b)$ maka :

$$P(a < X < b) = \int_a^b f(x) dx$$

Jika nilai probabilitas ini tinggi maka dapat diprediksikan bahwa observasi-observasi tersebut terakumulasi dalam interval tersebut. Fungsi densitas menginformasikan bahwa observasi-observasi lebih sering terjadi dalam interval tersebut.

Pada umumnya tidak bisa diketahui fungsi densitas secara langsung, sedangkan yang ada hanyalah sekumpulan data dari hasil observasi. Untuk mengatasi permasalahan ini dilakukan suatu pendekatan yaitu *pendekatan non parametrik*.

Estimasi densitas kernel merupakan metode yang cukup fleksibel dalam pendekatan fungsi densitas yang belum diketahui. Salah satu parameter yang sangat penting dalam estimasi densitas kernel adalah parameter penghalus yang menentukan derajat kehalusan estimator densitasnya.

Parameter penghalus ini dinamakan *bandwidth*. Perbedaan besar *bandwidth* akan menghasilkan bentuk kurva estimator yang berbeda-beda. Oleh karena itu, agar menghasilkan suatu estimator densitas yang mendekati fungsi densitas yang sebenarnya maka perlu pemilihan *bandwidth* yang optimal untuk estimator densitas kernel. Permasalahan yang ada adalah bagaimana cara pemilihan *bandwidth* sehingga diperoleh estimator densitas yang mendekati fungsi densitas yang sebenarnya. Salah satu metode yang digunakan untuk menentukan *bandwidth* optimal adalah metode Cross-validation. Komputasi dari metode ini dapat dilakukan dengan menggunakan dua macam algoritma yaitu :

1. Algoritma WARP (Weighted Averaging of Rounded Point).

Dalam algoritma digunakan fungsi pembobotan dan adanya pembagian data ke dalam bin-bin kecil. Akan tetapi algoritma ini tidak dibahas di sini.

2. Algoritma direct(langsung)

Dinamakan algoritma langsung karena data observasi langsung dimasukkan dalam perhitungan seleksi *bandwidth*.

Tujuan dari penulisan ini adalah :

1. Mengetahui cara pemilihan *bandwidth* optimal dengan metode Cross-validation.
2. Menerapkan metode Cross-validation pada suatu data untuk mendapatkan *bandwidth* optimal.

Pembahasan dalam tugas akhir ini akan dibatasi pada pemilihan bandwidth dengan metode Cross-validation dalam estimasi densitas kernel densitas kernel dari data-data pengamatan yang bervariasi tunggal dengan algoritma direct.

Bab I berisi tentang latar belakang, perumusan masalah dan sistematika mengenai tugas akhir ini.

Bab II berisi tentang teori penunjang yang meliputi pengertian fungsi densitas, estimasi fungsi densitas, estimator, kriteria kesalahan estimasi densitas serta estimator densitas kernel.

Bab III adalah materi inti yang memberikan penyelesaian mengenai seleksi bandwidth h dengan metode Cross-validation menggunakan algoritma direct.

Bab IV merupakan kesimpulan dari uraian yang telah dijelaskan pada bab sebelumnya.

