

PEMODELAN KELUHAN KESEHATAN DAN INDEKS KEBAHAGIAAN DI INDONESIA TAHUN 2021 MENGGUNAKAN PENDEKATAN *LOCAL POLYNOMIAL*

Ilham Faishal Mahdy

Statistika, Universitas Islam Bandung
Email Korespondensi: ilham.faishal@unisba.ac.id

ABSTRACT

To improve the well-being of the population based on community satisfaction, Statistics Indonesia (BPS) has been measuring the Happiness Index since 2012. Although the Indonesian social well-being index for 2021 is now higher than the previous 2017 measurement, the Covid-19 pandemic still hits communities today, causing health problems.. This study was conducted to examine how the rate of health complaints affects the happiness index in Indonesia in 2021 by using the Local Polynomial nonparametric regression model. The results show a fluctuating pattern that tends to decrease, which means that an increase in health complaints in the community tends to cause the health index to decrease.

Keywords: *Happiness Index, Health Complaints, Local Polynomial*

ABSTRAK

Dalam upaya peningkatan kesejahteraan penduduk yang berlandaskan pada kepuasan masyarakat, BPS melakukan pengukuran Indeks Kebahagiaan sejak tahun 2012. Saat ini Indeks Kebahagiaan masyarakat Indonesia di tahun 2021 meningkat dari pengukuran sebelumnya di tahun 2017, meskipun kondisi pandemi Covid-19 masih melanda hingga saat ini yang menimbulkan keluhan kesehatan pada masyarakat. Penelitian ini dilakukan untuk melihat bagaimana pengaruh persentase keluhan kesehatan terhadap Indeks Kebahagiaan di Indonesia pada tahun 2021 dengan menggunakan model regresi nonparametrik *Local Polynomial*. Hasil penelitian menunjukkan pola fluktuatif yang cenderung menurun, yang berarti peningkatan keluhan kesehatan di masyarakat cenderung menyebabkan indeks kesehatan menurun.

Kata kunci: *Indeks Kebahagiaan, Keluhan Kesehatan, Local Polynomial*

1. PENDAHULUAN

Indonesia merupakan salah satu negara terpadat di dunia dengan dengan populasi lebih dari 260 juta jiwa. Hal ini tentu saja menjadi tantangan pada pemerintah pusat dalam membangun negara dengan statusnya sebagai negara berkembang. Kualitas hidup masyarakat bisa menjadi tolak ukur bahwa suatu negara berhasil memberdayakan sumber daya manusianya.

Indikator kesejahteraan masyarakat tidak hanya dimaksudkan untuk menggambarkan keadaan kesejahteraan materi, tetapi juga keadaan kesejahteraan subjektif yang biasanya

didefinisikan sebagai keadaan mental yang baik, termasuk evaluasi hidup yang positif dan negatif serta tanggapan terhadap pengalaman tersebut. Dari perspektif ekonomi, kebahagiaan sulit untuk didefinisikan tetapi dapat diukur. Menurut (Ng, 1997), kebahagiaan didefinisikan sebagai kesejahteraan. Adapun (Easterlin, 1974) berpendapat bahwa kebahagiaan tidak lepas dari kesejahteraan subjektif, kepuasan, utilitas, dan kesejahteraan. (Frey & Stutzer, 2000) mendefinisikan kebahagiaan sebagai kesejahteraan subjektif. Semakin tinggi nilai indeksinya, maka semakin bahagia standar hidupnya. Sebaliknya, semakin rendah nilai indeksinya, maka semakin tidak bahagia.

Indikator kebahagiaan merupakan ukuran yang menggambarkan kesejahteraan, karena kebahagiaan mencerminkan kesejahteraan yang dicapai oleh setiap individu. Indikator kebahagiaan menggambarkan kesejahteraan subjektif di beberapa bidang kehidupan, yang dianggap perlu dan penting bagi sebagian besar penduduk dan masyarakat. Indeks kebahagiaan mengukur kebahagiaan atau kepuasan hidup penduduk pada skala 0-100. Indeks ini merupakan indeks gabungan yang terdiri dari tingkat kepuasan terhadap 10 aspek esensial kehidupan. Kesepuluh aspek tersebut mencerminkan tingkat kebahagiaan (Badan Pusat Statistik, 2015).

Di Indonesia, pengukuran tingkat kebahagiaan sudah dilakukan secara berkala setiap 3 tahun sekali. Dalam publikasi Indeks Kebahagiaan 2021 yang dirilis oleh Badan Pusat Statistik (Badan Pusat Statistik, 2021) dimuat ulasan mengenai hubungan antara Indeks Kebahagiaan dengan kemiskinan, pengangguran dan Indeks Pembangunan Manusia. Sedangkan keterkaitan antara kesehatan yang merupakan salah satu dari 10 aspek dalam kehidupan dengan indeks kebahagiaan belum dijelaskan dalam publikasi tersebut.

Keluhan kesehatan yang terjadi di penduduk Indonesia masih menjadi perhatian pada tahun 2021, terlebih jika melihat fenomena pandemi Covid-19 yang melanda di dunia, tidak terkecuali Indonesia. Mobilitas penduduk Indonesia yang cukup tinggi menyebabkan penyebaran Covid-19 di Indonesia terjadi dengan begitu cepat dan menimbulkan keluhan kesehatan di masyarakat. Sebagai negara dengan iklim tropis, Indonesia menghadapi beberapa penyakit endemik seperti demam berdarah, malaria dan tuberkulosis. Penyakit endemik terus berdampak luas terutama pada masyarakat di negara berkembang seperti Indonesia. Hal ini terkait dengan pembangunan yang tidak merata, kepadatan penduduk yang sulit diatur, hingga permasalahan ekonomi. Sehingga perlu dikaji bagaimana pengaruh keluhan kesehatan terhadap indeks kebahagiaan di masa pandemi Covid-19.

Dalam ilmu statistika, alat yang digunakan untuk mendeteksi pengaruh suatu variabel independen terhadap variabel dependen adalah analisis regresi. Seringkali kasus nyata yang dihadapi dalam penelitian adalah variabel berasal distribusi yang tidak diketahui. Sehingga diperlukan pendekatan regresi yang tidak biasa, yang disebut regresi nonparametrik. Pendekatan ini memiliki fleksibilitas yang tinggi (Eubank R. L., 1988). Hal ini didasarkan pada tidak adanya asumsi yang diperlukan dalam regresi nonparametrik. Adapun teknik yang dapat digunakan untuk mengestimasi parameter dalam regresi nonparametrik diantaranya dengan pemulusan menggunakan regresi kernel *Local Polynomial* dengan parameter pemulus (*bandwidth*) dipilih berdasarkan metode pembandingan.

2.1. Local Polynomial

Pendekatan lain dalam regresi kernel adalah *Local Polynomial*. Gagasan utama dari metode ini adalah melakukan aproksimasi secara lokal menggunakan derajat polinomial pada model regresi sebagai berikut:

$$y_i = m(x_i) + \varepsilon_i \quad (2.1)$$

dimana y_i adalah variabel dependen / respon pada pengamatan ke-i, m merupakan fungsi regresi yang tidak diketahui, x_i merupakan variabel independent / prediktor pada pengamatan ke-i, serta ε_i yang merupakan residual pada pengamatan ke-i. Aproksimasi ini didasari oleh prinsip deret Taylor yang menyatakan bahwa setiap fungsi mulus dapat diaproksimasi secara lokal dengan beberapa derajat polinomial. Bentuk dari model regresi polinomial adalah sebagai berikut (Wu & Zhang, 2006):

$$m(x_i) \approx m(x) + (x_i - x)m^{(1)}(x) + \dots + \frac{(x_i - x)^p m^{(p)}(x)}{p!} \quad (2.2)$$

dimana p merupakan derajat polinomial. Misalkan $\beta_r = m^{(r)}(X_i)/r!$ dengan r berada pada rentang 1 hingga p , maka persamaan (2.2) diatas dapat dituliskan sebagai berikut:

$$m(x_i) \approx \beta_0(x) + (x_i - x)\beta_1(x) + \dots + (x_i - x)^p \beta_p(x) \quad (2.3)$$

Persamaan (2.3) dapat dinyatakan sebagai berikut:

$$\mathbf{m}(x) = \mathbf{X}\boldsymbol{\beta} \quad (2.4)$$

dengan \mathbf{X} merupakan matriks berukuran $n \times (p + 1)$ dimana:

$$\mathbf{X} = \begin{bmatrix} 1 & (x_1 - x) & (x_1 - x)^2 & \dots & (x_1 - x)^p \\ 1 & (x_2 - x) & (x_2 - x)^2 & \dots & (x_2 - x)^p \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & (x_n - x) & (x_n - x)^2 & \dots & (x_n - x)^p \end{bmatrix}$$

$$\boldsymbol{\beta} = [\beta_0(x), \beta_1(x), \dots, \beta_p(x)]^t \text{ serta } \mathbf{m}(x) = [m(x_1), m(x_2), \dots, m(x_n)]^t$$

Untuk mendapatkan estimator $\hat{\beta}$ dilakukan dengan meminimumkan kriteria *Weighted Least Squares* (WLS) dengan menjadikan fungsi kernel sebagai pembobot (Chen, 2005). Adapun kriteria WLS dapat dituliskan sebagai berikut:

$$\sum_{i=1}^n (y_i - x_i\beta)^2 K_h(x_i - x) \quad (2.5)$$

dimana K_h merupakan fungsi kernel dimana $K_h(x) = \frac{1}{h} K\left(\frac{x}{h}\right)$ dengan $\int K(x)dx = 1$ serta memiliki *bandwidth* h . Berikut ini merupakan beberapa jenis fungsi Kernel yang dapat digunakan diantaranya Uniform, Segitiga, Epachnikov, Kuadrat, Twiweight, Cosinus, dan Gaussian (Härdle, 1990). Dalam penelitian ini, fungsi kernel yang akan digunakan adalah kernel Gaussian dengan rumus sebagai berikut:

$$K(x) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{1}{2}(-x^2)\right); -\infty < x < \infty$$

dengan x merupakan derajat kehalusan dari fungsi kernel dan $k = 0, \dots, x - 1$ merupakan urutan dari fungsi kernel, serta I merupakan indikator.

Adapun persamaan (2.5) dapat dituliskan dengan formula sebagai berikut:

$$S(\beta) = (y - X\beta)^t K_h (y - X\beta) \quad (2.6)$$

dimana $K_h = \text{diag}(K_h(x_1 - x), K_h(x_2 - x), \dots, K_h(x_n - x))$. Dengan demikian, penaksir koefisien regresi polinomial lokal $\hat{\beta}$ dapat dituliskan dengan formula sebagai berikut:

$$\hat{\beta} = (X^t K_h X)^{-1} X^t K_h y \quad (2.7)$$

2.2. Generalized Cross Validation (GCV)

Untuk mendapatkan model terbaik pada regresi kernel dan spline, maka model tersebut harus mengandung nilai *bandwidth* optimal. *Bandwidth* menentukan berapa banyak data yang digunakan agar sesuai dengan setiap *Local Polynomial*. Nilai dari *bandwidth* menentukan tingkat kemulusan dari fungsi regresi. Jika nilai *bandwidth* kecil maka fungsi regresi bersifat kasar yang menyebabkan nilai bias yang kecil dan nilai varians yang besar. Sebaliknya, jika nilai *bandwidth* besar maka fungsi regresi bersifat *smooth* yang menyebabkan nilai varians yang kecil serta nilai bias yang besar. Oleh karena itu, perlu dilakukan penentuan *bandwidth* optimal yang menghasilkan MSE terkecil. Salah satu metode untuk mendapatkan *bandwidth* yang optimal adalah metode GCV dengan rumus sebagai berikut (Golub, Heath, & Wahba, 1979):

$$GCV(h) = \frac{MSE(h)}{\left(\frac{1}{n} \text{tr}[I - A(\cdot)]\right)^2} \quad (2.5)$$

dimana $MSE(h) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$ serta I merupakan matriks identitas. Sedangkan $A(\cdot)$ diperoleh dari hubungan pada $\hat{y} = A(\cdot)y$ dimana $A(\cdot) = (X^t K_h X)^{-1} X^t K_h$. Nilai *bandwidth* yang optimum berkaitan dengan nilai GCV yang minimum (Eubank R. , 1999).

3. HASIL DAN PEMBAHASAN

3.1. Analisis Deskriptif

Berikut merupakan hasil analisis deskriptif pada variabel independen (X) dan variabel dependen (Y).

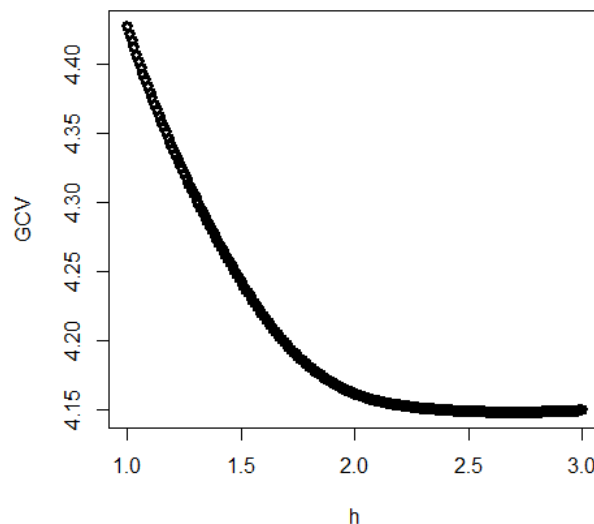
Tabel 1. Statistika Deskriptif

Variabel	Minimum	Median	Rata-Rata	Maksimum
X (Keluhan Kesehatan)	5.06	12.11	11.45	21.02
Y (Indeks Kebahagiaan)	68.08	73.07	72.74	76.34

Berdasarkan tabel 1 diperoleh informasi bahwa persentase penduduk yang mempunyai keluhan kesehatan pada tahun 2021 memiliki nilai maksimum sebesar 21.02, nilai minimum sebesar 5.06, median sebesar 12.11, serta rata-rata sebesar 12.11. Sedangkan indeks kebahagiaan pada tahun 2021 memiliki nilai maksimum sebesar 76.34, nilai minimum sebesar 68.08, median sebesar 73.07, serta rata-rata sebesar 72.74.

3.2. Pemilihan *Bandwidth*

Sebelum model regresi dibentuk, dilakukan simulasi untuk mencari *bandwidth* optimal dalam mengestimasi parameter regresi. Dalam hal ini, metode yang digunakan adalah GCV (*Generalized Cross Validation*). Dalam hal ini, *bandwidth* ditentukan dalam rentang [1,3] dengan selisih 0.01. Hasil dari perhitungan GCV untuk penentuan *bandwidth* optimal disajikan pada Gambar 2 berikut.

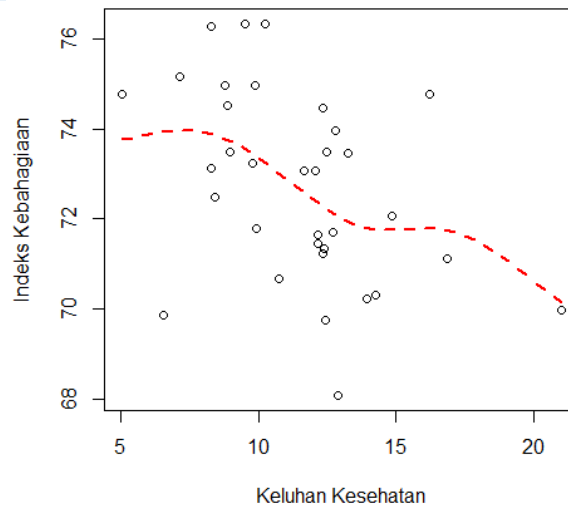


Gambar 2. Penentuan *Bandwidth* Optimal dengan Pendekatan GCV

Dengan menggunakan *software* R, diperoleh *bandwidth* optimal sebesar 2.7 yang memperoleh nilai GCV terkecil sebesar 4.148. Adapun derajat polinomial yang ditentukan adalah 2.

3.3. Polinomial Lokal

Berdasarkan *bandwidth* optimal yang diperoleh dari GCV, maka dilakukan estimasi parameter dengan menggunakan Polinomial Lokal. Adapun hasil estimasi dibandingkan dengan sebaran data disajikan pada Gambar 3 berikut.



Gambar 3. Hasil Estimasi Model *Local Polynomial*

Berdasarkan Gambar 3 terlihat bahwa kurva *Local Polynomial* menunjukkan pola fluktuatif yang cenderung menurun. Adapun ukuran kualitas dari model diukur dengan menggunakan RMSE. Dengan menggunakan *software* R, diperoleh informasi bahwa nilai RMSE dari model *Local Polynomial* tersebut adalah sebesar 2.675.

4. KESIMPULAN

Pemodelan nonparametrik pada data persentase penduduk yang mempunyai keluhan kesehatan dengan indeks kebahagiaan pada tahun 2021 menggunakan pendekatan *Local Polynomial* lokal dilakukan dengan fungsi kernel Gaussian menghasilkan pola fluktuatif yang cenderung menurun. Hal tersebut menunjukkan bahwa meningkatnya keluhan kesehatan yang dirasakan oleh penduduk cenderung akan mengakibatkan menurunnya indeks kebahagiaan.

5. DAFTAR PUSTAKA

- Badan Pusat Statistik. (2015). *Indeks Kebahagiaan Tahun 2014*. Jakarta: Badan Pusat Statistik.
- Badan Pusat Statistik. (2021). *Indeks Kebahagiaan 2021*. Jakarta: Badan Pusat Statistik.
- Badan Pusat Statistik. (2021). *Profil Statistik Kesehatan 2021*. Jakarta: Badan Pusat Statistik.
- Chen, B. (2005). Local Polynomial Regression Analysis of Clustered Data. *Biometrika*, 92 (1), 59-74.
- Easterlin, R. (1974). *Does Economic Growth Improvhe Human Lot?* In: Paul A. D., M. V. Reder (eds) *Nations and Households in Economic Growth: Essays in Honour of Moses Abramovitz*. New York:: Academic Press.
- Eubank, R. (1999). *Nonparametric Regression and Spline Smoothing*. New York: Marcel Dekker.
- Eubank, R. L. (1988). *Spline Smoothing and Nonparametric Regression*. New York: Merceel Dekker.
- Frey, B., & Stutzer, A. (2000). Happiness, economy and institutions. *Economic Journal Vol.110*, 918-938.
- Golub, G., Heath, M., & Wahba, G. (1979). Generalized Cross-Validation as a Method for Choosing a Good Ridge Parameter. *Technometrics*, 215-223.
- Hardle, W. (1990). *Applied Nonparametric Regression*. New York: Cambridge University.

- Ng, Y. (1997). A case for happiness, cardinal utility, and interpersonal comparability. *Economic Journal Vol.107 No.445*, 1848-1858.
- Purwanti, Y. (2022). Pengaruh Faktor Pendidikan dan Ekonomi pada Indeks Kebahagiaan di Indonesia. *Transformatif, IX(1)*, 1–13.
- Rositawati, A., & Budiantara, I. (2020). Pemodelan Indeks Kebahagiaan Provinsi di Indonesia Menggunakan Regresi Nonparametrik Spline Truncated. *Jurnal Sains dan Seni ITS, 8(2)*, 2337-3520.
- Takezawa, K. (2006). *Introduction to Nonparametric Regression*. New Jersey: John Wiley & Sons Inc.
- Wei, W. W. (1990). *Time Series Analysis: Univariate and Multivariate Methods*. Addison Wasley.
- Welsh, A., & Yee, T. (2005). Local Regression for Vector Responses. *Journal of Statistical Planning and Inference Vol. 136*, 3007-3031.
- Wu, H., & Zhang, J. (2006). *Nonparametric Regression Methods for Longitudinal Data Analysis*. New Jersey: John-Wiley and Sons Inc. Publication.