

**PENDUGAAN ANGKA PUTUS SEKOLAH DI KABUPATEN  
SEMARANG DENGAN METODE PREDIKSI TAK BIAS  
LINIER TERBAIK EMPIRIK PADA MODEL  
PENDUGAAN AREA KECIL**



**SKRIPSI**

**Disusun Oleh:  
NANDANG FAHMI JALALUDIN MALIK  
NIM. J2E 009 022**

**JURUSAN STATISTIKA  
FAKULTAS SAINS DAN MATEMATIKA  
UNIVERSITAS DIPONEGORO  
SEMARANG  
2013**

**PENDUGAAN ANGKA PUTUS SEKOLAH DI KABUPATEN SEMARANG  
DENGAN METODE PREDIKSI TAK BIAS LINIER TERBAIK EMPIRIK  
PADA MODEL PENDUGAAN AREA KECIL**

**Disusun Oleh :**

**NANDANG FAHMI JALALUDIN MALIK**

**J2E 009 022**

Skripsi

Sebagai Salah Satu Syarat untuk Memperoleh Gelar  
Sarjana Sains pada Jurusan Statistika

JURUSAN STATISTIKA  
FAKULTAS SAINS DAN MATEMATIKA  
UNIVERSITAS DIPONEGORO  
SEMARANG

2013

## HALAMAN PENGESAHAN I

Judul : Pendugaan Angka Putus Sekolah di Kabupaten Semarang dengan Metode Prediksi Tak Bias Linier Terbaik Empirik pada Model Pendugaan Area Kecil.  
Nama : Nandang Fahmi Jalaludin Malik  
NIM : J2E 009 022  
Jurusan : Statistika

Telah diujikan pada sidang Tugas Akhir tanggal 16 Oktober 2013 dan dinyatakan lulus pada tanggal 30 Oktober 2013

Semarang, 30 Oktober 2013

Mengetahui,  
Ketua Jurusan Statistika  
Fakultas Sains dan Matematika UNDIP,



**Dra. Hj. Dwi Ispriyanti, M.Si**  
NIP. 1957 09 14 1986 03 2 001

Panitia Penguji Ujian Tugas Akhir  
Ketua,

**Prof. Drs. H. Mustafid, M.Eng, Ph.D**  
NIP. 1955 05 28 1980 03 1 002

## HALAMAN PENGESAHAN II

Judul : Pendugaan Angka Putus Sekolah di Kabupaten Semarang dengan Metode Prediksi Tak Bias Linier Terbaik Empirik pada Model Pendugaan Area Kecil.

Nama : Nandang Fahmi Jalaludin Malik

NIM : J2E 009 022

Jurusan : Statistika

telah diujikan pada sidang Tugas Akhir tanggal 16 Oktober 2013

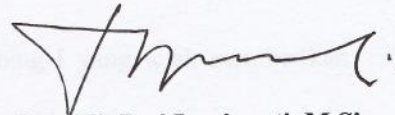
Semarang, 16 Oktober 2013

Pembimbing I



**Abdul Hoyyi, S.Si, M.Si**  
NIP. 1972 02 02 2008 01 1 018

Pembimbing II



**Dra. Hj. Dwi Ispriyanti, M.Si**  
NIP. 1957 09 14 1986 03 2 001

## KATA PENGANTAR

Puji syukur kehadirat Allah SWT, karena atas rahmat, hidayah, kemudahan, dan segala limpahan nikmat-Nya sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul “**Pendugaan Angka Putus Sekolah di Kabupaten Semarang dengan Metode Prediksi Takbias Linier Terbaik Empirik pada Model Pendugaan Area Kecil.**”

Penulis menyadari tanpa adanya dukungan dan bantuan dari berbagai pihak maka penulisan tugas akhir ini tidak akan berjalan dengan baik. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan terima kasih kepada :

1. Dra. Hj. Dwi Ispriyanti, M.Si selaku Ketua Jurusan Statistika FSM Universitas Diponegoro dan Dosen pembimbing II yang telah memberikan bimbingan serta pengarahan dalam penulisan Tugas Akhir ini.
2. Abdul Hoyyi, S.Si, M.Si selaku dosen pembimbing I yang telah memberikan bimbingan dan pengarahan dalam penulisan Tugas Akhir ini.
3. Bapak dan Ibu Dosen Jurusan Statistika, FSM Universitas Diponegoro yang telah memberikan ilmu yang sangat berguna.
4. Semua pihak yang telah memberikan bantuan, semangat dan doa sehingga penulis dapat menyelesaikan penulisan Tugas Akhir ini.

Penulis sadar bahwa Tugas Akhir ini masih jauh dari sempurna, untuk itu penulis mengharapkan kritik dan saran untuk perbaikan lebih lanjut.

Semarang, Oktober 2013

**Penulis**

## Abstrak

Pada saat ini informasi pada area kecil yaitu area yang memiliki ukuran contoh kecil sangat dibutuhkan. Pendugaan secara langsung pada area kecil akan menghasilkan nilai ragam yang besar. Oleh karena itu, diperlukan alternatif lain yang bisa digunakan yaitu dengan pendugaan tidak langsung. Pendugaan area kecil merupakan suatu metode pendugaan tidak langsung yang digunakan untuk menduga parameter pada area kecil dengan memanfaatkan informasi dari luar area, dari dalam area itu sendiri, dan dari luar survey. Salah satu metode yang sering digunakan adalah metode Prediksi takbias linier terbaik empirik (EBLUP). EBLUP ini akan digunakan untuk menduga angka putus sekolah tiap desa di Kabupaten Semarang. Informasi tambahan yang digunakan dalam metode EBLUP ini yaitu jumlah sarana pendidikan, jumlah penduduk, rata-rata pengeluaran perkapita dan jarak dari desa ke Kabupaten. Dari hasil pendugaan EBLUP memperlihatkan bahwa desa yang paling rendah angka putus sekolahnya adalah desa Beji dan yang paling tinggi adalah desa Pledokan. Pendugaan tidak langsung dengan metode EBLUP untuk kasus angka putus sekolah di Kabupaten Semarang memiliki koefisien variansi yang lebih kecil 0,598% daripada koefisien variansi yang didapat dari pendugaan langsung.

**Kata Kunci:** Angka Putus Sekolah (APts), Pendugaan Area Kecil (SAE), Prediksi takbias linier terbaik empirik (EBLUP)

## Abstract

Nowadays, small area information that has a small sample size is needed. A direct estimation in the small area will produce a large variance of values. In order of that, another alternative is needed that can be used is the indirect estimation. Small area estimation is an indirect estimation method that can be used to estimate parameters in a small area by utilizing information from outside the area, from the area itself, and from outside the survey. One of the methods that can be used is the empirical best linear unbiased prediction (EBLUP). EBLUP will be used to estimate the dropout rate for each village in the district of Semarang. Additional information used in this EBLUP method are the number of educational facilities, population, average expenditure per capita and distance from village to district. The results of EBLUP estimation showed that the lowest dropout rate village is Beji village and the highest is Pledokan village. Indirect estimation with EBLUP methods for the case of dropout rate in the district of Semarang has a coefficient variance 0,598% smaller than the coefficient variance that obtained from direct estimation.

**Keywords :** Dropout Rate, Small Area Estimation (SAE), Empirical Best Linear Unbiased Prediction (EBLUP)

## DAFTAR ISI

	halaman
HALAMAN JUDUL .....	i
PENGESAHAN I .....	ii
PENGESAHAN II .....	iii
KATA PENGANTAR .....	iv
ABSTRAK .....	v
ABSTRACT .....	vi
DAFTAR ISI .....	vii
DAFTAR SIMBOL .....	x
DAFTAR TABEL .....	xii
DAFTAR GAMBAR .....	xiii
DAFTAR LAMPIRAN .....	xiv
<b>BAB I PENDAHULUAN</b>	
1.1 Latar Belakang .....	1
1.2 Tujuan Penulisan .....	3
<b>BAB II TINJAUAN PUSTAKA</b>	
2.1 Angka Putus Sekolah .....	4
2.1.1 Definisi Angka Putus Sekolah .....	4
2.1.2 Definisi Anak Putus Sekolah .....	4
2.1.3 Alasan Tidak Sekolah .....	6
2.1.4 Hak Akan Pendidikan .....	7

2.2 Profil Kabupaten Semarang.....	7
2.3 Pengertian Penduga dan Sifat Penduga .....	9
2.4 Uji Asumsi Normalitas .....	10
2.5 Korelasi <i>Pearson Product Moment</i> .....	11
2.6 Metode Maksimum <i>Likelihood</i> .....	13
2.7 Penduga Langsung ( <i>Direct Estimation</i> ) .....	14
2.8 Penduga Tidak Langsung ( <i>Indirect Estimation</i> ).....	14
2.8.1 Pendugaan Area Kecil ( <i>Small Area Estimation</i> ) .....	14
2.8.1.1 Model Area Kecil .....	16
2.9.2 Metode <i>Empirical Best Linear Unbiased Prediction</i> (EBLUP) .....	17

### **BAB III METODOLOGI PENELITIAN**

3.1 Sumber Data .....	21
3.2 Variabel Penelitian .....	21
3.3 Langkah-langkah Penelitian .....	22
3.4 Diagram Alir Analisis Data .....	23

### **BAB IV HASIL DAN PEMBAHASAN**

4.1 Pendugaan Langsung Angka Putus Sekolah .....	24
4.2 Pendugaan Tidak Langsung.....	26
4.2.1 Pemilihan Variabel Pendukung .....	26

4.2.2 Model <i>Small Area Estimation</i> dengan Metode <i>Empirical Best Linear Unbiased Prediction</i> (EBLUP).....	29
--	----

**BAB V PENUTUP**

5.1 Kesimpulan.....	34
5.2 Saran .....	34

<b>DAFTAR PUSTAKA .....</b>	<b>35</b>
-----------------------------	-----------

<b>LAMPIRAN .....</b>	<b>37</b>
-----------------------	-----------

## DAFTAR SIMBOL

$f(x)$	: Fungsi densitas peluang dari variabel acak $X$
$F(x)$	: Fungsi distribusi kumulatif dari variabel acak $X$
$E(X)$	: Nilai harapan dari variabel acak $X$
$v_i$	: Pengaruh acak area kecil
$e_i$	: Error sampling
$\beta$	: Koefisien regresi (pengaruh tetap)
$b_i$	: Konstanta positif
$r$	: Koefisien korelasi
$\pi$	: Pi
$L$	: Likelihood
$l$	: Log-likelihood
$d$	: Turunan
$n$	: Ukuran sampel
$dk$	: Derajat bebas
iid	: <i>Independent and identically distributed</i>
$\mathbf{X}_{n \times p}$	: Matriks ukuran $n \times p$ atau matriks variabel pendukung
$y$	: Vektor ukuran $n \times 1$ atau matriks variabel respon
$\mathbf{Z}_{n \times q}$	: Matriks ukuran $n \times q$ yang hanya mengandung intersep
$N$	: Distribusi normal
$\sim$	: Berdistribusi
$\mathbf{G}$	: Matriks hasil perkalian antara varians pengaruh acak dengan matriks identitas

**R** : Matriks hasil perkalian antara varians error sampling dengan matriks identitas

$\hat{\theta}_i^{BLUP}$  : Parameter penduga BLUP

$\gamma$  : Gamma

V : Varians dari variabel respon (y)

Ln : Logaritma natural

$\mathfrak{T}$  : Fraktur T

$\hat{\theta}_i^{EBLUP}$  : Parameter penduga EBLUP

## DAFTAR TABEL

	halaman
Tabel 2.1 Tabel Interpretasi Koefisien Korelasi Nilai r .....	12
Tabel 4.1 Nilai Statistik Angka Putus Sekolah Hasil Pendugaan Langsung ...	24
Tabel 4.2 Nilai Korelasi <i>Pearson</i> .....	28
Tabel 4.3 Nilai Dugaan Parameter Beta.....	30
Tabel 4.4 Nilai Statistik Angka Putus Sekolah Hasil Pendugaan EBLUP .....	32

## DAFTAR GAMBAR

	halaman
4.1 Diagram <i>Boxplot</i> Angka Putus Sekolah Hasil Pendugaan Langsung .....	25
4.2 Histogram Angka Putus Sekolah Hasil Pendugaan Langsung.....	26
4.3 Diagram Pencar Variabel Respon dengan Variabel Pendukung.....	27
4.4 Histogram Pengaruh Acak ( <i>Efek Random</i> ) .....	29

## DAFTAR LAMPIRAN

	halaman
Lampiran 1 Data Angka Putus Sekolah Pendugaan Langsung.....	37
Lampiran 2 Uji Normalitas .....	39
Lampiran 3 Output Nilai Korelasi <i>Pearson</i> .....	42
Lampiran 4 Data Angka Putus Sekolah Pendugaan Tidak Langsung.....	43
Lampiran 5 Program EBLUP Menggunakan Software SAS .....	45
Lampiran 6 Output Program SAS .....	48
Lampiran 7 Data Variabel Pendukung.....	50
Lampiran 8 Tabel Nilai-Nilai Kritis Uji <i>Kolmogorov Smirnov</i> .....	52

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Dalam perkembangan statistika banyak metode untuk mendapatkan data, diantaranya sensus dan survey. Sensus maupun survey berperan penting dalam proses pengambilan keputusan yang berbasis pada data. Survey sendiri sering dilakukan secara rutin baik di lembaga penelitian swasta maupun negeri. Tujuan utama dari survey adalah untuk mendapatkan informasi mengenai parameter populasi dengan mengefektifkan biaya yang tersedia. Selama ini survey rutin yang dilakukan oleh pemerintahan suatu negara hanya dirancang untuk memperoleh informasi data skala nasional. Persoalan muncul ketika dari survey seperti ini ingin diperoleh informasi untuk area yang lebih kecil, seperti informasi level provinsi, Kabupaten/kota bahkan mungkin level kecamatan dan desa/kelurahan.

Dalam konteks survey, penduga dikatakan langsung (*direct estimator*) apabila pendugaan terhadap parameter populasi pada wilayah hanya didasarkan terhadap data contoh yang diperoleh dari wilayah tersebut. Pendugaan langsung umumnya didasarkan pada teknik penarikan contohnya (*sampling technique*). Misalnya, *simple random sampling*, *stratified random sampling*, *cluster sampling*, dan sebagainya. Pendugaan secara langsung pada area kecil akan menghasilkan nilai varians yang besar jika contoh yang diambil berasal dari survey yang dirancang untuk skala besar/nasional. Hal ini disebabkan oleh ukuran contoh yang diambil pada area tersebut kecil. Salah satu solusi yang digunakan adalah melakukan pendugaan tidak langsung dengan cara menambahkan variabel-

variabel pendukung dalam menduga parameter. Variabel pendukung tersebut berupa informasi dari area lain yang serupa, survey terdahulu pada area yang sama, atau variabel lain yang berhubungan dengan variabel yang ingin diduga. Pendugaan tidak langsung tersebut dikenal sebagai pendugaan area kecil atau lebih dikenal dengan *Small Area Estimation* (SAE).

Berbagai metode pendugaan area kecil (*small area estimation*) telah dikembangkan khususnya menyangkut metode yang berbasis model (*model-based estimator*). Berkembangnya otonomi daerah di Indonesia semakin membutuhkan statistik area kecil. Setiap pemerintahan daerah memiliki wewenang lebih dalam memajukan daerahnya. Kebutuhan statistik area kecil pada level Kabupaten/kota, kecamatan ataupun desa/kelurahan sangat penting sebagai dasar pemerintah daerah untuk menyusun sistem perencanaan, pemantauan dan penilaian pembangunan daerah atau kebijakan penting lainnya.

Dari penjelasan di atas, maka penulis dalam hal ini akan mencoba mengaplikasikan metode *small area estimation* dalam menduga angka putus sekolah ditingkat Desa/Kelurahan yang ada di Kabupaten Semarang. Menurut Badan Pusat Statistik angka putus sekolah dihitung berdasarkan jumlah penduduk pada kelompok umur 7 sampai 12 tahun, 13 sampai 15 tahun dan 16 sampai 18 tahun yang tidak bersekolah lagi. Dalam mengaplikasikan metode tersebut, untuk menduga angka putus sekolah dengan berdasarkan kelompok umur diatas, penulis mengalami kesulitan dalam mendapatkan data yang akan menjadi variabel pendukungnya. Oleh karena itu, perhitungan angka putus sekolah pada tugas akhir ini tidak berdasarkan kelompok umur 7 sampai 12 tahun, 13 sampai 15 tahun dan

16 sampai 18 tahun, tetapi berdasarkan pada kelompok umur secara keseluruhan yaitu 7 sampai 18 tahun.

Parameter yang menjadi perhatian atau variabel responnya diambil dari hasil Survey Sosial Ekonomi Nasional (SUSENAS) tahun 2012, dan variabel pendukung diambil dari hasil sensus Potensi Desa (PODES) tahun 2012 yang dilaksanakan oleh Badan Pusat Statistik Provinsi Jawa Tengah. Permasalahan yang akan dibahas yaitu bagaimana mendapatkan hasil penduga tidak langsung dari angka putus sekolah dengan menggunakan salah satu metode dalam pendugaan area kecil yaitu prediksi takbias linier terbaik empirik atau biasa disebut dengan *Empirical Best Linear Unbiased Prediction* (EBLUP).

## **1.2 Tujuan Penulisan**

Tujuan dari penulisan tugas akhir ini yaitu mendapatkan pendugaan angka putus sekolah disetiap desa/kelurahan yang ada di Kabupaten Semarang berdasarkan metode *Empirical Best Linear Unbiased Prediction* (EBLUP).