

- REGRESI LINIER PERPUSTAKAAN UNIVERSITAS AIRLANGGA

- LINEAR MODEL (STATISTIK)

PEMILIHAN VARIABEL PADA REGRESI LINIER DENGAN METODE STATISTIK C_p MALLOWS

SKRIPSI

MPM 43/04

Zak

P



MILIK
PERPUSTAKAAN
UNIVERSITAS AIRLANGGA
SURABAYA

PURNADI ZAKARIA

JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS AIRLANGGA
SURABAYA
2003

PEMILIHAN VARIABEL PADA REGRESI LINIER DENGAN METODE STATISTIK C_p MALLOWS

SKRIPSI

Sebagai Salah Satu Syarat untuk Memperoleh
Gelar Sarjana Sains Bidang Matematika pada Fakultas Matematika dan
Ilmu Pengetahuan Alam Universitas Airlangga.



Pembimbing I,

Ir. Elly Ana, M.Si
NIP 131 837 441

Pembimbing II,

Drs. H. Sediono, M.Si
NIP 131 653 448

LEMBAR PENGESAHAN SKRIPSI

Judul : Pemilihan Variabel pada Regresi Linier dengan Metode Statistik

Cp Mallows

Nama : PURNADI ZAKARIA

NIM : 089711592

Tanggal Ujian : 26 September 2003

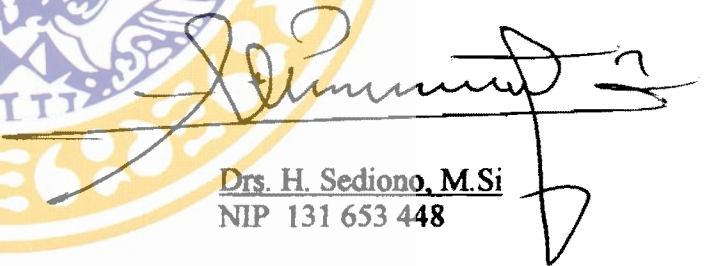
Disetujui Oleh :

Pembimbing I,

Pembimbing II,



Ir. Elly Ana, M.Si
NIP 131 837 441



Drs. H. Sediono, M.Si
NIP 131 653 448

Dekan Fakultas MIPA
Universitas Airlangga



Drs. H. A. Latief Burhan, MS
NIP 131 286 709

Ketua Jurusan Matematika
FMIPA UNAIR



Drs. MOH. Imam Utoyo, M.Si
NIP 131 801 397

Purnadi Zakaria, 2003. Pemilihan variabel pada regresi linier dengan metode statistik C_p Mallows. Skripsi ini dibawah bimbingan Ir. Elly Ana, M.Si dan Drs. H. Sediono, M.Si. Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Airlangga

ABSTRAK

Tulisan ini bertujuan untuk mengetahui keidentikan statistik T_p yang didefinisikan $T_p = W_p - K + 2p$ dengan statistik C_p Mallows yang didefinisikan $C_p = \frac{JKS_p}{MSE} - n + 2p$ dengan K menyatakan jumlah parameter pada model penuh dan p menyatakan jumlah parameter pada submodel (model yang telah disederhanakan). Pada estimasi kuadrat terkecil, kedua metode ini akan identik dengan nilai $C_p \leq p$ dan nilai $T_p \leq p$.

Kelebihan C_p Mallows dapat dibuktikan dengan aplikasi pada data Physical Fitness Course yang dilakukan D.R. A.C. Courtesy Linnerud yang menggunakan 6 variabel bebas. Kemudian diseleksi dengan metode semua kemungkinan regresi dengan kriteria nilai $C_p \leq p$ sehingga diperoleh persamaan regresi terbaik yang melibatkan 4 variabel bebas yaitu :

$$\hat{Y} = 98,15 - 0,2X_1 - 2,77X_3 + 3,5X_5 + 0,27X_6$$

Kata-kata kunci : *regresi linier, C_p Mallows, statistik T_p*

Purnadi Zakaria, 2003. Pemilihan variabel pada regresi linier dengan metode statistik C_p Mallows. The advisors this thesis are Elly Ana, Ir, M.Si dan H. Sediono, Drs, M.Si. Mathematics Departement and Natural Science Faculty Airlangga University

ABSTRACT

The purpose of this paper is to know the identical of statistic T_p which $T_p = W_p - K + 2p$ and statistic C_p Mallows is $C_p = \frac{JKS_p}{MSE} - n + 2p$ and K is total parameter of full model and p is total parameter of submodel (simple model). On estimation of least square, both of these model will identical with $C_p \leq p$ and $T_p \leq p$.

The advantage of C_p Mallows can be proved with the application at data Physical Fitness Course that was done by D.R. A.C. Courtesy Linnerud using free variables 6. Afterwards, it was select by all possible regression method with $C_p \leq p$ criteria until was obtained the best regression similarity that involve free variables 4, i.e :

$$\hat{Y} = 98,15 - 0,2X_1 - 2,77X_3 - 3,5X_5 + 0,27X_6$$

Key Words : linear regression, C_p Mallows, statistic of T_p