E0 259
DATA ANALYTICS

ASSIGNMENT-1

Duckworth Lewis Stern Method

By

PAVAN KUMAR S V

19629

MTech Aerospace Engineering

OBJECTIVE:

The task assigned in this assignment is to fit an objective function of the form

$$Z(u,w) = Z_0(w)[1 - \exp\{-Lu/Z_0(w)\}]$$

Where,
u is the number of overs,
w number of wickets,
Z_0 is the average which depends on the number of wickets in hand,
L is the slope of the graph.

The above function is a run production function which takes the wickets in hand and number of overs remaining as input and gives the Predicted run as output, this method is useful for Predicting run in case the cricket match is not completely played due to rain or some other situation.

The error function used for this objective is the sum of squared errors loss function which as to minimized in order to determine the parameters Z_0 and L.

After fitting the model 10 plots have to be plotted with the obtained 10 Z_0 values and one L value.

APPROACH USED:

1. The data file given contains data of cricket matches where both the data for 1st innings and 2nd innings are given out of which the second innings data is not required for fitting the run predictor function. So, we eliminate the second innings data by using the pandas library in python.
2. Now in the given data some of the matches have not been played for whole 50 overs due to some reasons such data are deleted because taking such data into account makes approximations bit high, so I deleted such data by using pandas library.
3. The given over data is overs used which I converted to overs remaining by subtracting 50.
4. With all the above-mentioned Data preprocessing a Data set with overs remaining, runs scored and wickets in hand as been generated.
5. Now by using the SciPy optimize library the objective function that is the least square error loss function is minimized.

6. The least square error loss function is

$$Error = (1/n)\sum_{i=1}^{n}(y_i - Z(u,w)_i)^2$$

7. The above function is minimized by using the SciPy optimize minimize function with 'BFGS' method and the final parameters and loss value was obtained after optimizing the loss function.

RESULT:

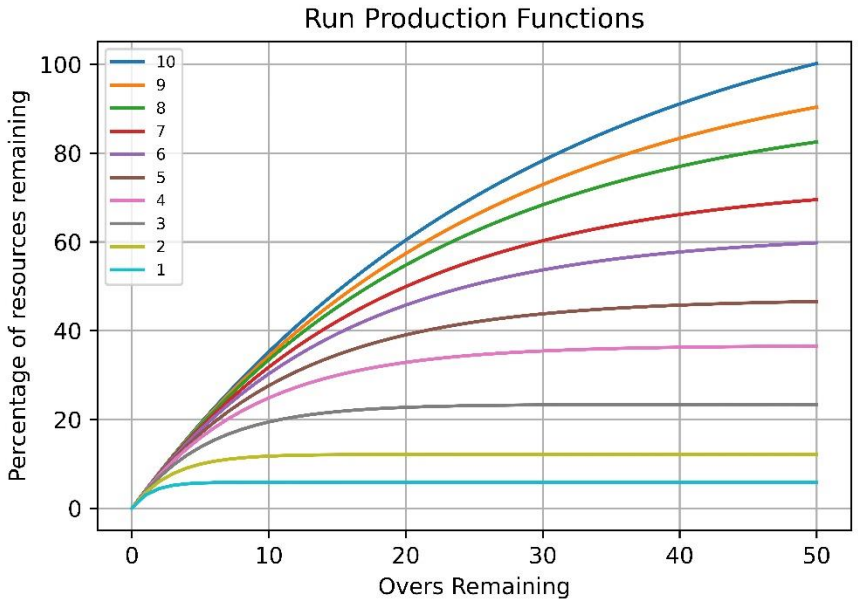
The Eleven parameters obtained after optimizing the loss function is

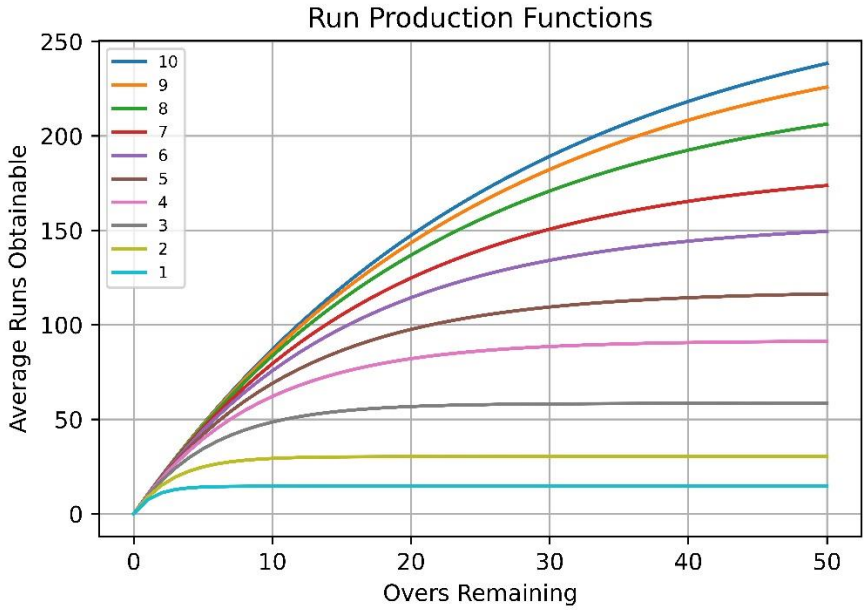
$Z_0(10)$	307.03762384
$Z_0(9)$	261.94118901
$Z_0(8)$	230.44984607
$Z_0(7)$	184.9234788
$Z_0(6)$	154.88108798
$Z_0(5)$	117.82047911
$Z_0(4)$	91.59632526
$Z_0(3)$	58.4580264
$Z_0(2)$	30.2971033
$Z_0(1)$	14.58928328
L	10.38756537

The Total loss with normalization = 1360.398143633372

The Total loss without normalization = 79917949.34588607

PLOT:


Run Production Functions
Percentage of resources remaining
Overs Remaining
10
9
8
7
6
5
4
3
2
1
0
20
40
60
80
100
0
10
20
30
40
50


Run Production Functions
Average Runs Obtainable
Overs Remaining
10
9
8
7
6
5
4
3
2
1
0
50
100
150
200
250
0
10
20
30
40
50