

# Application of Discriminant Analysis in Forecasting



# Purpose of Discriminant Analysis

Classify the subjects into mutually exclusive classes or groups on the basis of some information on the predictor variables (indep/disc).

To attain the goal, a linear combination of the predictor variables are formed.

This linear combination is used to classify the subjects.

The classification is to be done in such a manner that the within group variance is as small as possible.

# Understanding Purpose of Discriminant Analysis

## Let's Frame Two Examples

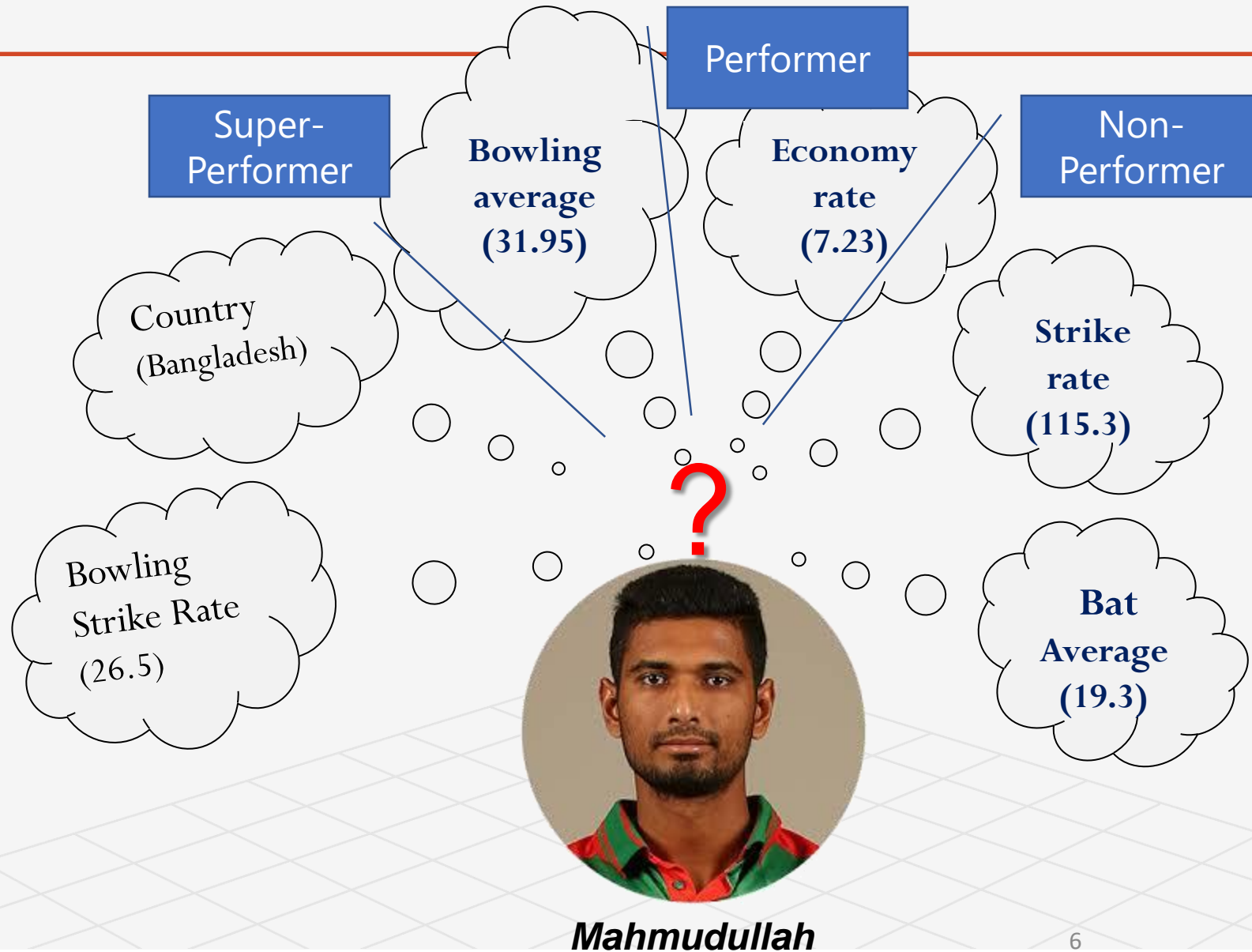
- 1. Bidding for  
an all-rounder  
in next IPL  
Auction**
- 2. Deciding  
about a loan  
by the banker**

---

## Bidding for an all-rounder in next IPL Auction

Name	Country	BSR ( $X_1$ )	ER ( $X_2$ )	BA ( $X_3$ )	Avg. ( $X_4$ )	SR ( $X_5$ )	Outcome
SP Narain	West Indies	20.06	6.02	22.53	13..8	126.72	Super Perf
DJ Bravo	West Indies	20.04	8.46	29.85	21	116.41	Super Perf
SA Hasan	Banglades h	18.23	6.8	20.66	23.08	121.33	Performer
K Pollard	West Indies	21.13	8.26	29.09	20.21	132.19	Non-Perf
S Raina	India	26.08	7.57	32.92	29.7	132.96	Performer
R Jadeja	India	26.16	7.27	31.71	9.67	93.55	Performer
A Russell	West Indies	24.33	9.24	37.48	14.38	129.21	Non-perf
JP Dumini	S Africa	22.16	7.75	28.63	36.96	123.9	

# A new allrounder is offered for bidding



	Name	Identity Check	Address Proof	Credit Score	Income / month	A/C in bank	Salary A/C	Job type	Loan Decision	
	Anawar	1	1	578	78243	1	1	Salaried 1	Yes	
	Binit	1	1	630	64534	1	0	Salaried 1	Yes	
	Chadni	1	0	732	21010	0	0	Salaried 2	No	
	David	1	0	686	103476	0	0	Business	Yes	
	Juliya	1	1	816	76345	1	0	Business	No	
	June	1	1	800	122673	1	1	Salaried 1	Yes	
	Manas	1	1	706	87435	0	0	Salaried 1	No	
	Nandu	1	0	712	102360	0	0	Salaried	Yes	

# Will I get the loan?

Identity: OK

Address: OK

Credit Score:  
675

**Income: Rs,  
60000**

A/C in Bank: No

**Salaried A/C :  
No**

Job type: Business





# Assumptions

The observations are random in nature and the independent variables are normally distributed.

There must be at least two groups and each case belongs to only one groups.

Groups shall be defined before collecting the data.

The var-covar matrix of the groups shall be identical

# How Discriminant Analysis works?

Based on the newly defined axes each observation is transformed using a function called as the discriminating function, D

Namely,  $D = v_1X_1 + v_2X_2 + \dots + v_pX_p + a$

$v_i$  is the discriminant coefficient (weight) corresponding to the  $i^{\text{th}}$  variable.

'a' is a constant.

Discriminating Variable

Grouping  
Variable

Training  
sample

Name	Identity Check	Address Proof	Credit Score	Income/ month	A/C in bank	Salary A/C	Job type	Loan Decision
Anawar	1	1	578	78243	1	1	Salaried1	Yes
Binit	1	1	630	64534	1	0	Salaried1	Yes
Chadni	1	0	732	21010	0	0	Salaried2	No
David	1	0	686	103476	0	0	Business	Yes
Juliya	1	1	816	76345	1	0	Business	No
June	1	1	800	122673	1	1	Salaried1	Yes
Manas	1	1	706	87435	0	0	Salaried1	No
Nandu	1	0	712	192369	0	0	Salaried2	Yes
Opu	1	1	590	117640	1	0	Business	Yes
Salim	1	1	670	25731	0	0	Salaried2	No

For each observation, the value of the discriminant function i.e.  $D$  is determined.

In case of a two-group classification, the observations with  $D$  value less than a given cut-off value are classified to one group

---

While those with larger cut-off values are classified into another group.

In case of classification with more than two groups ( $k$ , say) there shall be  $k - 1$  discriminant function.

$$D = 0.2 \text{ Iden} + .3 \text{ Add} + 0.27 \text{ CS} - 1.3 \text{ Inc} + \dots - 1.1 \text{ JT} + 3.8$$

Identity: OK

Address: OK

Credit Score:  
675

Income: Rs,  
60000

A/C in Bank: No

Salaried A/C :  
No

Job type: Business

$$D = 7.85$$

10

