Analysis of IJSO questions

P K Joshi
HBCSE-TIFR
MUMBAI
Traditional way of calculating DI

- **Discrimination Index**: for any given MCQ
- How well does a question filter out “good” performing students from rest
- Does the question help us in selecting good students from the lot?

- List out the answers of each student: “correct” or “not-correct”
- Sort the answers on the basis of “total marks”; decreasing order of total marks obtained
Sort in decreasing order on total marks column CD
Number of students in top 27% list, who answered question 5 correctly; \( N_t \)
Number of students in bottom 27% list, who answered question 5 correctly; \( N_b \)

\[
DI = \frac{N_t - N_b}{C} \quad \text{with} \quad C = 0.27 \times \text{number of students}
\]
Extension of the analysis

• Presented last year
• Handle negative numbers (negative marking)
• Handle long questions (theory & experiments)
• How to re-define “total” on which the sorting is carried out.
• Computer code to calculate in traditional style
• Computer code to calculate the extended version
• It can sort on all possible combinations of remaining questions to give DI as a number “range“ rather than a single number.
IJSO – 2013

30 MCQs
5 Theory questions
5 practical/experiment questions

How easy is a question?
Difficulty index Diff-I

Diff-I = Sum of all marks obtained in Question/(N x max marks)
Q2. Visual question on moon
IJSO-2017

The diagram shows a scatter plot with the DI (possibly a performance metric) on the y-axis and Q number on the x-axis. The data points are indicated with error bars. The plot includes horizontal lines at different values: Excellent (1.0), good (0.8), acceptable (0.4), and poor (0.2). The values of DI vary across the Q numbers, with some points clustered around the performance categories.
Conclusions:

Data available for 3 years

The question-by-question analysis can be carried out by future hosts

Questions with high Diff-I (very easy) and low DI values can be identified

Future hosting countries can look at such questions to avoid framing such questions

The pattern of high DI value questions can also be identified

Some high Diff-I questions are desirable?
Thank you