Report of the Committee of Statisticians, constituted by Kerala Public Service Commission to study the standardization procedure for Degree Level Common Preliminary Examinations, 2021.

The Kerala Public Service Commission constituted the committee to standardize the marks obtained by candidates in degree level common preliminary examinations conducted in two phases with different question papers based on the same syllabus. The committee has the opinion that the standardization is required only if there is significant variation in difficulty levels in between the different question papers.

One cannot determine which question is more difficult simply by reading the questions. It is not fair to decide a question is more difficult based on intuition or subjectivity of a person. The decision can be taken only based on empirical evidence.

So, the committee decided to compute the index of difficulty (difficulty level) of each question in each question paper. **The index of difficulty (p) of a question is defined as the proportion of correct answers of that question (that is the number of correct answers of the question divided by the total number of candidates who wrote the examination with the respective question paper) (Nitko (1996), Crocker & Algina, 1986).** The larger the proportion getting a question right, the easier the question. The higher the difficulty index indicates the easier the question and the lower the index of difficulty indicates the question is difficult. The index of difficulty will be always in between 0 and 1. **The index of difficulty '0' means the maximum is the difficulty and '1' means the difficulty is minimum.** There are two practices in the literature for the computation of difficulty indices. 1) based on the upper 27% and the lower 27% groups and 2) based on the total candidates. Several studies shown that, there will be no significant difference in difficulty indices based on two procedures when the size of the data is large. Here also the indices based on two procedures computed and confirmed that there is no significant variation between them. **Thus, we suggest the computation of difficulty indices based on total candidates.**

The questions in each question papers are to be divided into five strata as follows:

Stratum No.	Range of difficulty index(p)	Difficulty Level
1	$0 \leq p < 0.2$	Very difficult
2	$0.2 \leq p < 0.4$	Difficult
3	$0.4 \leq p < 0.6$	Average
4	$0.6 \leq p < 0.8$	Easy
5	$0.8 \leq p \leq 1$	Very easy

All questions in a question paper are to be stratified into five levels as given above and found the number of questions in each stratum out of 100 questions. It is denoted as below.

Stratum No.	Number of questions	
	Phase 1	Phase 2
1	N_{11}	N_{21}
2	N_{12}	N_{22}
3	N_{13}	N_{23}
4	N_{14}	N_{24}
5	N_{15}	N_{25}

If the distribution of difficulty levels in different question papers varied significantly, then we can infer that all question papers are not in the same level. Under this circumstance we have to adopt some methods to standardize the marks.

The Committee considered different procedures for standardization and illustrated with sample data. No procedure will be absolutely correct. The Committee was of the opinion that the effect of standardization should be same for all candidates who scored the same mark within a particular group (answered a particular question paper) because the weightage of all questions were same while conducting the examinations. Also it is not fair to give the

same benefits of difficulty for all the candidates in a stream but to be given according to their performance in the examination. **The benefit of standardization is given in proportion to the performance of the candidates in the respective question paper after equated the difficulty score of that question paper to the difficulty score of the question paper which is observed as minimum difficult by the above said procedure. Thus, if a candidate scored a mark zero or below zero, no benefit of standardization will be obtained.**

The committee unanimously suggests the following procedure for standardization.

A score of difficulty (DS) is to be computed for each question paper as given below.

Compute $DS_i = \sum[(N_{i1} - N_{i5})x0.9 + (N_{i2} - N_{i4})x0.7 + N_{i3}x0.5] / 100$

Where DS_i is the Score of difficulty for the ith question paper; N_{ij} is the number of questions in the jth stratum of the ith question paper. The weights of the Very difficult, Difficult and Average strata are the median scores of range of difficulty(p), 0.9, 0.7 and 0.5 respectively.

Then the proportion of difficulty (Multiplier, K_i) of each question paper to the lowest difficult question paper (DS_i is minimum) can be computed as:

$K_i = (DS_i / DS_{min})$

DS_{min} is the score of difficulty of the question paper which has the minimum DS.

Then the final mark of the jth candidate who attended the ith question paper (S_{ij}) is to be computed as:

$S_{ij} = Minimum(M_{ij} \times K_i, 100)$

Where M_{ij} is the mark actually scored **out of 100** (including negative marks) by the jth candidate in the ith group.

===

References

1. Crocker L. and Algina J. (1986). Introduction to classical and modern test theory, New York: Holt, Rinehart and Winston.

2. Nitko A.J (1996): Educational assessment of students. Second edition, New Jersey, USA, Prentice- Hall.