Z-Score Normalization

About Normalization

Normalization means adjusting values measured on different scales to a notionally common scale

Need for Normalization in Exam

Exam pertaining for a particular post/course could be spread across multiple shifts which will have different question paper for each shift. Hence the normalization of scores need to be carried out for all the candidates who had written the exam, across shifts for the same post/course.

Score Normalization logic

For each shift, average and standard deviation is calculated. Average tells us about overall scoring. Standard deviation tells about spread of marks. Take example below of one shift and see how average and StdDev is calculated.

Sr.	X (raw score)	x = (X-Xav)	x^2
1	31	-34	1156
2	46	-19	361
3	40	-25	625

PC 2/4/2019

4	71	6	36
5	65	0	0
6	90	25	625
7	59	-6	36
8	84	19	361
9	99	34	1156
N = 9	Total = L = ∑X = 585		$\sum x^2 = 4356$

Xav = : L / No. of candidates present for that particular shift = 585/9 = 65

Standard Deviation = $\sqrt{(\sum x^2)/N} = \sqrt{4356/9} = 484 = 22$

As shown in example above for one shift, same process is followed for every shift to calculate mean and StdDev of each shift.

Choosing the Base Shift

Then to calculate normalized score for each candidate, first we need to consider one shift as base shift. Statistically any shift can be considered as base shift because normalization will put all candidates properly in relative position.

PC 2/4/2019

But to avoid reduction in marks after normalization, we normally consider shift with highest mean as base shift. Also, in particular shift, if number of candidates is too low, we may not consider that shift as base shift (for statistical purposes) even if it has highest mean. (Generally, this threshold is kept at 70%).

Normalized Score for each candidate (Xn)

To calculate normalized score of each candidate, following formula is used.

$$Xn = \frac{S2}{S1} * (X - Xav) + Yav$$

S2	Is the Standard Deviation of the base shift
S1	Standard Deviation for the candidate shift (to be scaled to S2)
X	Raw score of a candidate
Xav	Simple average of the candidate Shift
Yav	Average of the base shift

PC 2/4/2019