SES'S L.S.RAHEJA COLLEGE OF ARTS AND COMMERCE

Course: Testing And Assessment Unit: IV of Semester VI Prepared by: Ms. Puja Patwardhan

UNIT 4

STANDARD SCORES

- A raw score that has been converted from one scale to another scale, the latter scale has some arbitrarily set mean & SD.
- Reasons→

1. More easily interpretable.
2. The position of a testtaker's performance relative to other testtakers is readily apparent.

- Different systems for standard score- each unique in terms of its respective mean & SDs.

z Score-

- A standard scale that may be thought of as the zero plus or minus one scale.
- Has a mean set at 0 & SD set at 1.
- ***Results from the conversion of a raw score into a number indicating how many SD units the raw score is below or above the mean***.
- Eg- National spelling test- a raw score of 65
- Z= (X-mean)/sd
- Value can be positive or negative.
- A z score is equal to the difference between a particular raw score & the mean divided by the SD.
- Eg- a raw score= 65, mean= 50, SD= 15
- a raw score of 65 was found to be equal to a z score of +1.
- Knowing about a z score of 1 on a spelling test provides context and meaning for the score.

- Normal curve→ only about 16% of the other testtakers obtained higher scores.
- Knowing only the raw score does not provide any useful information→ context lacking.
- Standard scores provide a convenient context for comparing scores on different tests.
- ABC's raw score on the hypothetical Main Street Reading Test=24 & Arithmetic Test=42.
- Without knowing anything other than these raw scores→ ABC did better on the arithmetic test.
- 2 z scores would be more informative.
- Z score based on the performance of other students in the class, reading test= 1.32, arithmetic= -0.75.
- Raw score in arithmetic was higher; z score provides a different picture.
- Z scores indicate→ relative to other students in the class, ABC performed above average on the reading test and below average on the arithmetic.
- An interpretation of exactly how much better→ obtained by reference to tables detailing distances under the normal curve + the resulting percentage of cases that could be expected to fall above or below.

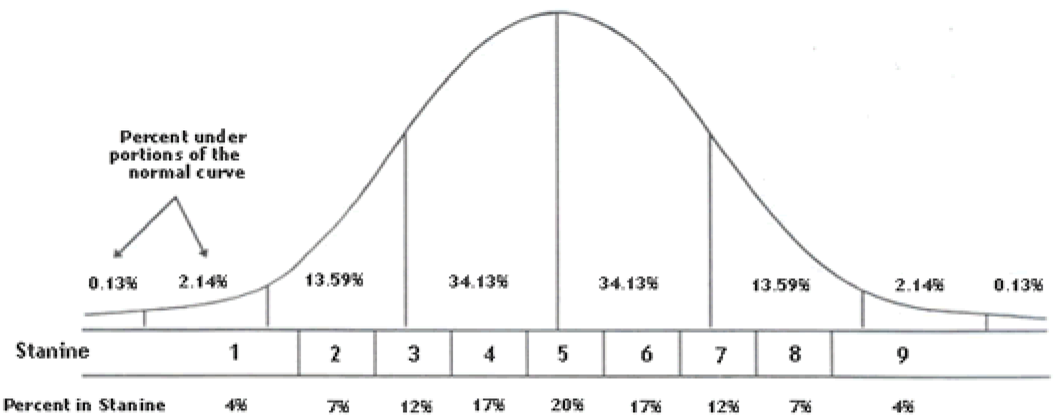
T scores-

- The scale used in the computation of T scores→ 50 plus or minus 10 scale.
- A scale with a mean set at 50 and a SD at 10.
- Devised by McCall (1922, 39), named a T score in honor of his professor Thorndike.
- T score is composed of a scale that ranges from 5 SDs below the mean & 5 SDs above.
- Eg- a raw score that fell exactly at 5 SDs below the mean would be equal to a T score of 0, at the mean=50 and 5 SDs above=100.
- ***Advantage***- none of the scores is negative→ can make further computation cumbersome.

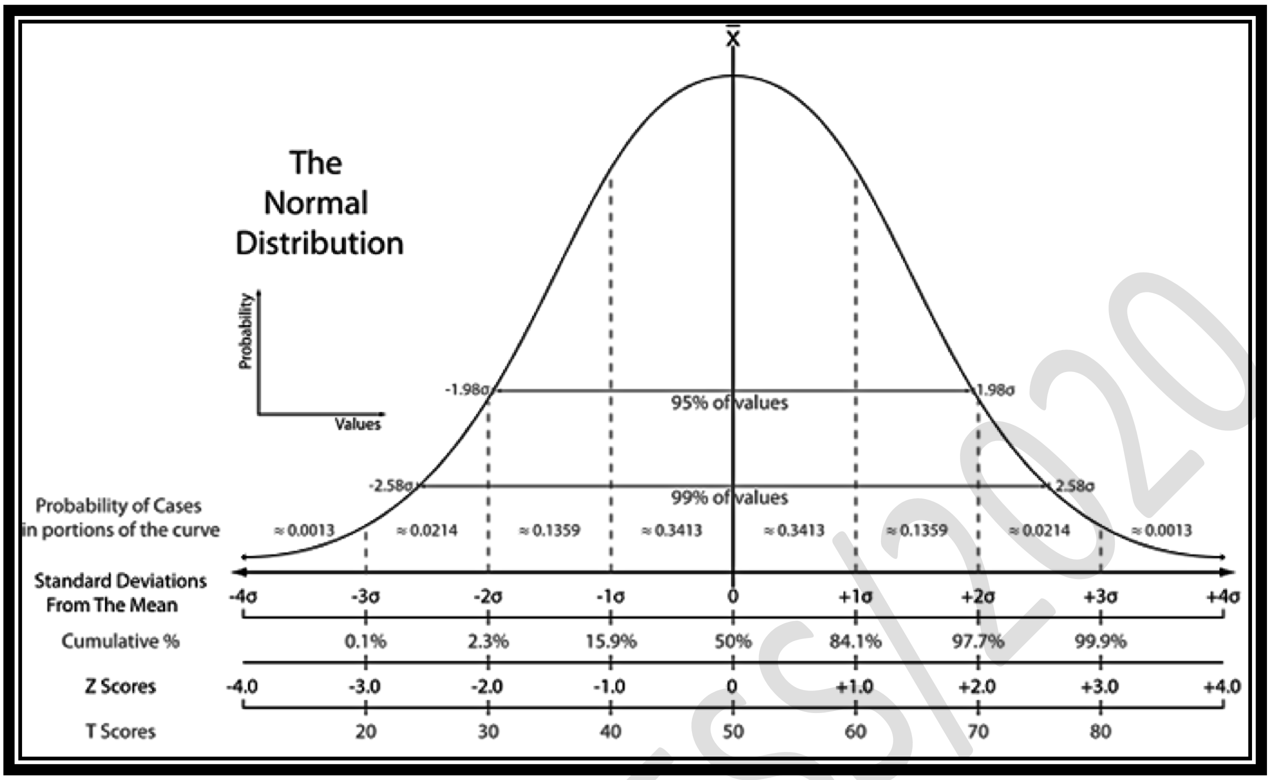
STANINE- *standard-nine*

Figure 2

The dimensions contained in the evaluation are measured by the Stanine System. This system of measurement incorporates a line broken into nine standad sections. The nine Standardized sections can be compared to the Bell Curve of the general population, as shown in the illustration below. A score in the 1-3 stanine range equates to the bottom one-third of the population on the curve; a score in the 4-6 range equates to the mid-range of the population on the curve, and a score in the 7-9 range equates to the upper one-third of the population on the curve



- Developed during WW-2→ standard score with a mean of 5 & SD of approximately 2.
- Commonly used in achievement tests.
- Stanines are different from other standard scores→ take on whole values from 1-9
- Represent a range of performance that is half of a SD in width.
- The 5th stanine indicates performance in the average range, from 1/4 SD below the mean to ¼ SD above→ captures the middle 20% of the scores in a normal distribution.
- The 4th & 6th stanines→ 1/2 SD wide and capture the 17 % of cases below & above the 5th stanine.



Linear transformation-

- Scores converted from raw scores may involve either linear or nonlinear transformations.
- A standard score obtained by a linear transformation→ retains a direct numerical relationship to the original raw score.
- The magnitude of differences between such standard scores exactly parallels the difference between corresponding raw scores.
- Sometimes scores may undergo more than one transformation.
- Eg- the creators of the SAT- 2nd linear transformation on their data to convert z scores into a new scale that has a mean=500 & SD=100

Nonlinear transformation-

- Required when the data under consideration are not normally distributed→ comparisons with normal distribution need to be made.
- The resulting standard score does not necessarily have a direct numerical relationship to the original raw score.
- The original distribution should be normalized.

NORMALIZED STANDARD SCORES-

- Conceptually, normalizing a distribution involves 'stretching' the skewed curve into a shape of a normal curve and creating a corresponding scale of standard scores, a scale that is technically referred to as a **normalized standard score scale.**
- Normalization of a skewed distribution of scores is desirable for purposes of comparability. One of the primary advantages of a standard score on a test is that it can be readily compared with a standard score on another test.
- However, such comparisons are appropriate only when the distributions from which they derived are the same. In most cases they are the same when they are normal.
- But if, for e.g. distribution A were normal and distribution B were highly skewed, then *z* scores in these respective distributions would represent different amounts of area subsumed under the curve. A *z* score of -1in distribution A tells us, about 84% of the scores fall above this score. A *z* score in distribution B mean for e.g. only 62% of the scores were higher than that score.
- It is generally preferable to fine tune the test according to difficulty or other relevant variables so that the resulting distribution will approximate the normal curve. This is so because there are technical cautions to be observed before attempting normalization.