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

Probability and Normal Probability Curve

PROBABILITY:

Probability is the measure of the likelihood that an event will happen. It is quantified as a number between 0 and 1 (where 0 indicates impossibility and 1 indicates certainty). The higher the probability of an event, the more certain we are that the event will occur.

A simple example is the tossing of an unbiased coin. Since the coin is unbiased, the two outcomes (head or tails) are equally probable. Since no other outcome is possible, the probability is ½ or 50% of either 'head' or 'tail'.

LAWS OF PROBABILITY:

1. ***Probability of success + probability of failure= 1***

- Probability of success (p)= ½
- Probability of failure (q)= ½
- ***p+q=1***

2. Rule of addition-

The probability of occurrence of any one of several particular events is the sum of their individual probabilities, provided they are mutually exclusive.

- Eg→ tossing a coin→ we want to get a head, the probability of that occurrence is ½
- After tossing a coin, we get a head (success)→ probability of success in this case is ½
- If we get a tail (failure), probability of failure is ½
- Thus, probability of success + probability of failure=1
- The rule applies the use of the word '***OR***'
- Sometimes addition theorem is also known as 'OR' rule.

LSRC/TUT-LES/2020

- The rule is valid only when the outcomes are mutually exclusive→ occurrence of one precludes the occurrence of any of the others.

3. Rule of Subtraction

Rule of Subtraction the probability that event will occur is equal to 1 minus the probability that event A will not occur.

$$P(A) = 1 - P(A')$$

Suppose, for example, the probability that Bill will graduate from college is 0.80. What is the probability that Bill will not graduate from college? Based on the rule of subtraction, the probability that Bill will not graduate is 1.00 - 0.80 or 0.20.

4. Rule of Multiplication:

The rule of multiplication applies to the situation when we want to know the probability of the intersection of two events; that is, we want to know the probability that two events (Event A and Event B) both occur.

Rule of Multiplication The probability that Events A and B both occur is equal to the probability that Event A occurs times the probability that Event B occurs, given that A has occurred.

$$P(A \cap B) = P(A)\, P(B|A)$$

Example

An urn contains 6 red marbles and 4 black marbles. Two marbles are drawn without replacement from the urn. What is the probability that both of the marbles are black?

Solution: Let A = the event that the first marble is black; and let B = the event that the second marble is black. We know the following:

In the beginning, there are 10 marbles in the urn, 4 of which are black. Therefore, $P(A) = 4/10$.

After the first selection, there are 9 marbles in the urn, 3 of which are black. Therefore, $P(B|A) = 3/9$.

Therefore, based on the rule of multiplication:

$$P(A \cap B) = P(A)\, P(B|A)$$

$$P(A \cap B) = (4/10) * (3/9) = 12/90 = 2/15$$

LSRC/TUT-LESS/2020

NORMAL PROBABILITY CURVE:

The literal meaning of the term normal is average. Most of the things like intelligence, wealth, beauty, height etc. are quite equally distributed. There are quite a few persons who deviate noticeably from average, either above or below it. If we plot such a distribution on a graph paper, we get a bell-shaped curve, referred to as *Normal Curve.*

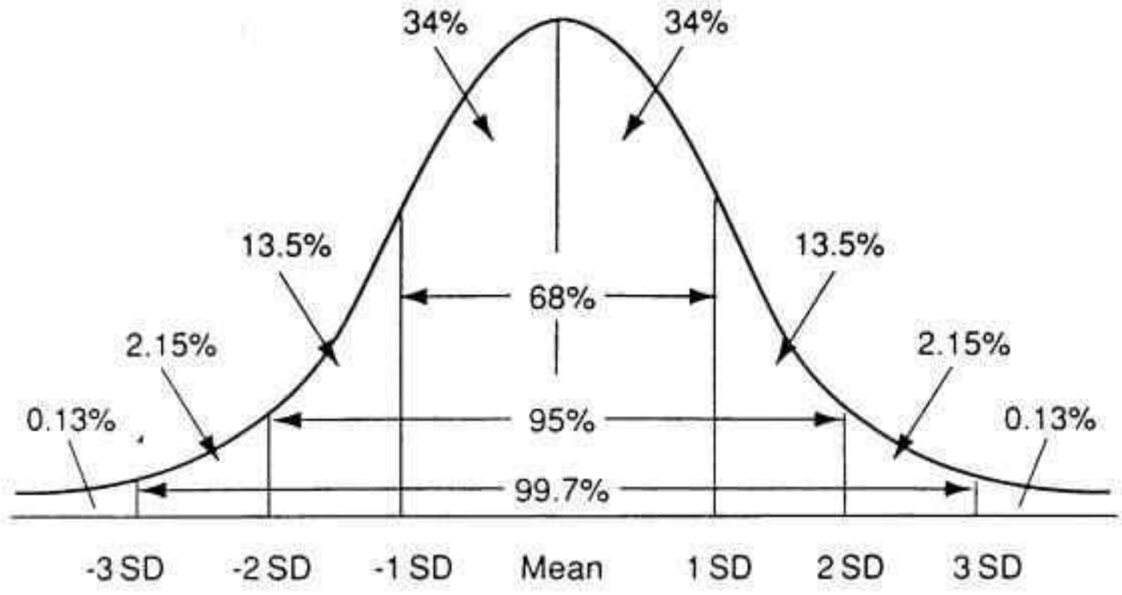
The data from a certain coin or a dice throwing experiment involving a chance success or probability; if plotted on a graph paper gives a frequency curve which closely resembles the normal curve. Hence, it is also known as *Normal Probability Curve.*

Normal curve was derived by Laplace and Gauss (1777-1855) independently. They also named it ‘curve of error’, where ‘error’ is used in the sense of a deviation from the normal, true value. In the honour of Gauss, it’s also known as *Gaussian Curve’*.

The normal curve takes into account the law which states that the greater the deviation from the mean or an average, the less frequently it occurs. For e.g. in terms of Intelligence, its rare to find people with very low or very high intelligence. It’s normally distributed in the population.

LSRC/TUT-LES 2020

AREA UNDER NORMAL CURVE:



1. 50% of the scores occur above the mean and 50% below.
2. Approximately 34% between the mean and 1 SD above mean.
3. Approximately 34% between the mean and 1 SD below mean.
4. Approximately 68% of all scores occur between the mean & +/-1SD
5. Approximately 95% between the mean and +/-2 SDs.
6. Approximately 99% of the scores fall between -3 and +3 SDs.
7. The area on the normal curve between 2 and 3 SDs above & 2 and 3 SDs below the mean are known as tails.
8. The normal curve has 2 tails.

CHARACETRISTICS OF NORMAL CURVE:

a) For this curve, mean, median and mode are the same.
b) The curve is perfectly symmetrical. In the sense, it is not skewed. The value of measured skewness for normal curve is zero.
c) The normal curve serves as a model for describing the flatness or peakedness of a curve through the measure of kurtosis. For the normal curve, the value of kurtosis is 0.263.

LSRC/TUT-LESSON

d) The curve is asymptotic. It approaches but never touches the X-axis. It is because of the possibility of locating in the population a case which scores still higher than the highest score or still lower than the lowest score. Therefore, theoretically, it extends from minus infinity to plus infinity.

e) As the curve does not touch the base line, the mean is used as the starting point for working with the normal curve.

f) To find out deviations from the mean, standard deviation of the distribution (σ) is used as a unit of measurement.

g) The curve extends on both sides -3σ distance on the left to $+3\sigma$ on the right.

APPLICATIONS AND IMPORTANCE OF NORMAL CURVE:

a. Used as a model- Normal curve represents a model distribution. It can be used as a model to -
   i. Compare various distributions with it i.e. to say whether the distribution is normal or not and if not, in what way it diverges from the normal
   ii. Compare two or more distributions in terms of overlapping; and
   iii. Evaluate students' performance from their scores.

b. Computing Percentiles and Percentile Ranks- Normal probability curve may be conveniently used for computing percentiles and percentile ranks in a given normal distribution.

c. Applying the concept of standard error of measurement- The normal curve is also known as the normal curve of error or simply the curve of error on the grounds that it helps in understanding the concept of standard error of measurement.

d. Ability Grouping- A group of individuals may be conveniently grouped into certain categories as A, B, C, D and E (very good, good, average, poor, very poor) in terms of some trait with the help of a normal curve.

e. Transforming and combining qualitative data- Under the assumption of normality of the distributed variable, the sets of qualitative data such as ratings, letter grades and categorical ranks on a scale may be conveniently transformed and combined to provide an average rating for each individual.

f. Conversion into comparable standard scores- With the help of the normal curve, we can convert the raw scores belonging to different tests into standard normalized scores like sigma and T-scores.

LSRC/TUT-LESS/2020

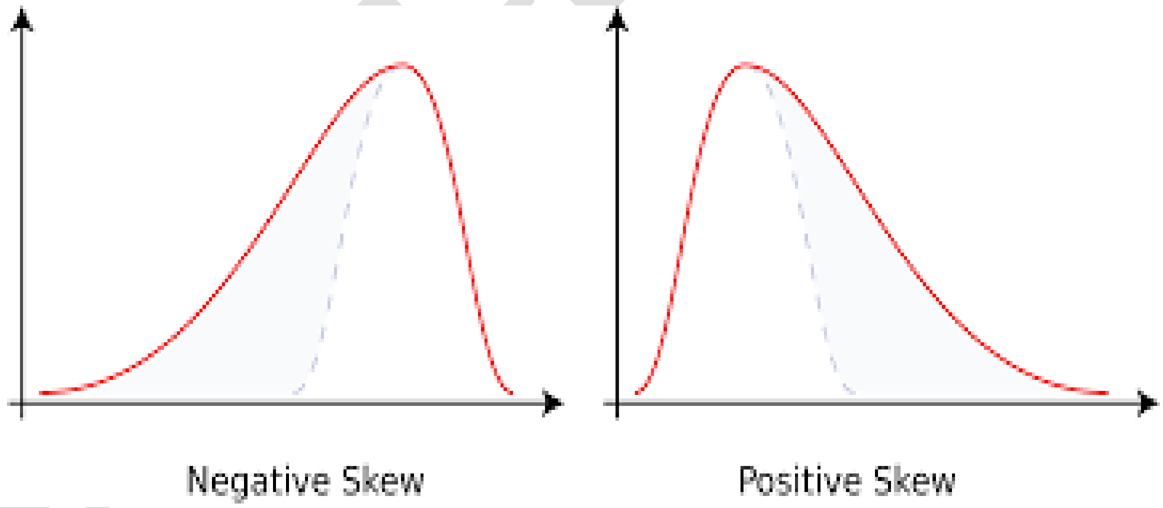
g. Determining the relative difficulty of test items- Normal curve provides the simplest rational method of scaling test items for difficulty and therefore, may be conveniently employed for determining the relative difficulty of test questions, problems and other test items.

- **Importance**-

1. Knowledge of the areas can be useful to the interpreter of test data.
2. Tells us about where the score falls among a distribution of scores
3. Tells about a person & people who share that person's life.
4. This knowledge conveys about how impressive, average or lackluster the individual is with respect to a particular ability.
5. Conveys useful information about a test score in relation to other test scores.

SKEWNESS-

- Distribution can be characterized by their **skewness**→ ***the nature and extent to which symmetry is absent***.
- It is an indication of how the measurements in a distribution are distributed.



- A distribution is skewed when the mean & the median fall at different points in the distribution.
- The balance is shifted to one side-> lacks symmetry.
- In a normal distribution→ skewness 0→ mean & median are equal/ at the same point.
- ***Positive skew***→ when ***relatively few of the scores fall at the high end*** of the distribution.

- Eg- positively skewed examination results may indicate that the test was too difficult.

- More items that were easier would have been desirable in order to better discriminate at the lower end.
- ***Negative skew***→ relatively ***few scores fall at the low end*** of the distribution.
- Results may indicate that the test was too easy.
- More items of a higher difficulty level would make it possible to better discriminate between scores at the upper end.
- Skewness→ carries a ***negative implication***.
- Associated with ***abnormal***→ because the skewed distribution deviates from the symmetrical/ normal distribution.
- The presence/ absence of symmetry in a distribution is simply one characteristic by which a distribution can be described.
- Eg- a hypothetical Marine Corps Ability & Endurance Screening Test→ administered to al civilians seeking to enlist in the Marines.
- Which figure would best describe the data?
- A level of difficulty would have been built into the test to ensure that relatively few assesses would score at the high end.
- Most of the applicants→ score at the low end.
- Consistent with the objective→ looking only for a few good men.
- Skewness is not→ good/ bad/ abnormal
- Various **formulas** for measuring skewness-
- One way→ Examination of relative distances of quartiles from the median.
- In a **positively skewed** distribution Q3-Q2 will be **greater** than the distance of Q2-Q1.
- **Negatively** skewed→ Q3-Q2 will **be less** than the distance of Q2-Q1.
- **Symmetrical** distribution→ the distances from Q1 and Q3 to the median are the **same**.
- Skewness, ***Sk= 3 (mean- median)/ SD***

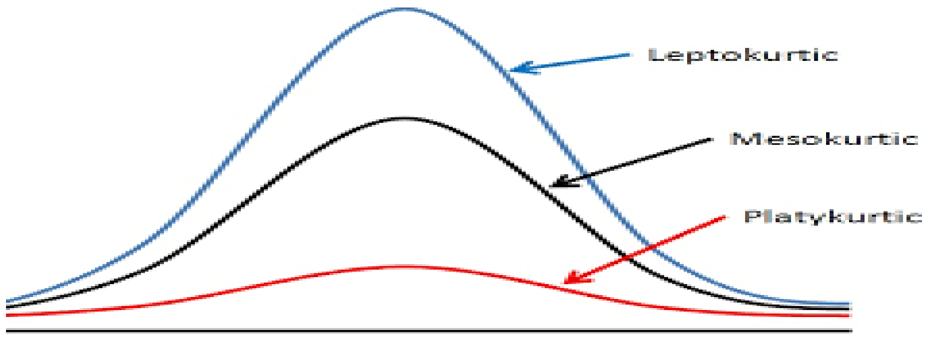
LSRC/TUT-LESS/2020

- Skewness will be 0 when the mean and median coincide.
- Skewness will be positive when the value of the mean is higher than the median.
- Negative→ value of mean is less than the median.

KURTOSIS-

- Steepness of a distribution is in its center.
- Root kurtic is added to one of the prefixes platy-, lepto- or meso- to describe the peakedness/ flatness of 3 general types of curves.

1. ***Platykurtic***→ relatively flat
2. ***Leptokurtic***→ relatively peaked
3. ***Mesokurtic***→ somewhere in the middle



- Many methods exist for measuring kurtosis.
- Computer programs feature an index of skewness that ranges from -3 to +3.
- Technical matters related to the measurement and interpretation of kurtosis are controversial among specialists.
- A simple measure for estimating kurtosis in terms of percentiles is,

$$\mathbf{Ku = Q/P_{90}-P_{10}}$$

Q= quartile deviation

P= percentile

- For the normal curve (mesokurtic distribution) the value of Ku= 0.263

LSRC/IC-SS/2020

- If the value of kurtosis is less than 0.263→ leptokurtic distribution.
- Value more than 0.263→ platykurtic.

Causes of skewness & kurtosis-

1. ***Selection***→ biased sample.
2. ***Unsuitable or poorly made tests***→ normality depends upon the no. of items & their difficulty level.

- Test is too easy or too difficult.

3. ***Non-normal distribution***→ lack of normality in the trait being measured.

- Eg- death rate would be maximum in the early ages and the rate would decrease with increase in age.
- Distribution would be positively skewed.

4. ***Errors in the use of tests***→ administration & scoring

LSRC/TUT-LESS/2020

LSRC/TUT-LESS/2020

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.

LSRC/TUT-LESS/2020

- 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-

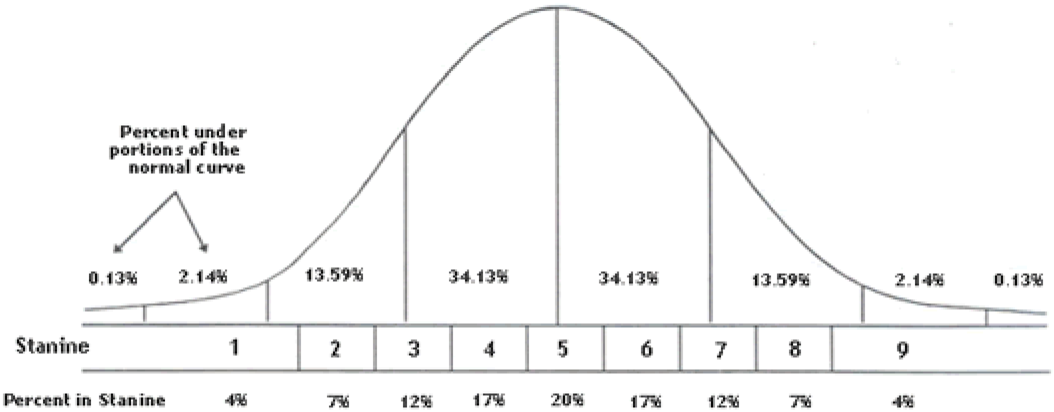
- The scale used in the computation of T scores→ 50 plus or minus 10 scale.
- A scale with a mean set at 50 and a SD at 10.
- Devised by McCall (1922, 39), named a T score in honor of his professor Thorndike.
- T score is composed of a scale that ranges from 5 SDs below the mean & 5 SDs above.
- Eg- 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*

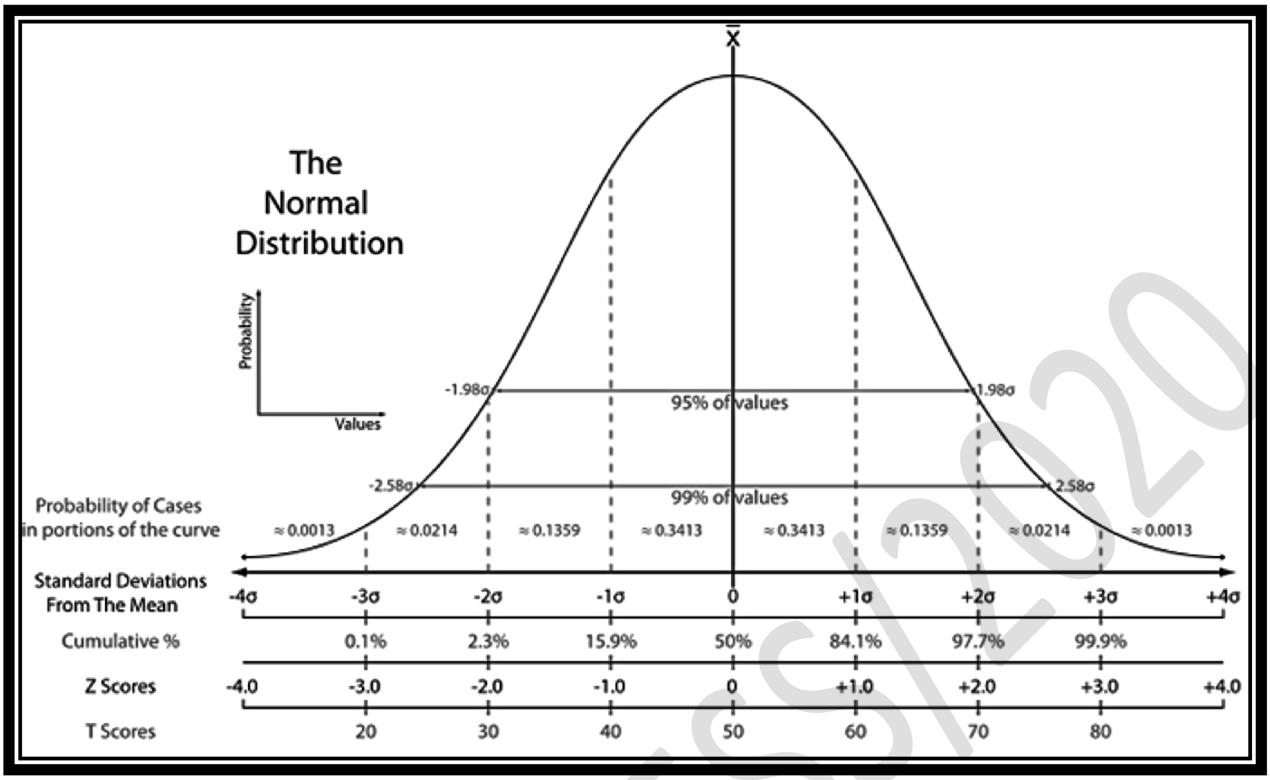
LSRC/TUT-LESS/2020

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.

LSRC/TUT-LESS/2020

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.

LSRC/TUT-LESS/2020

LSRC/TUT-LESS/2020