1. Introduction

Welcome to the series of E-learning modules on illustrations on Tchebyscheff's and Markov's inequalities. In this module we are going cover the concept of Tchebyscheff's and Markov's inequalities and the procedure to apply these inequalities to practical problems.

By the end of this session, you will be able to:

- Describe Tchebyscheff's and Markov's inequalities
- Explain the procedure to apply Tchebyscheff's inequality
- Apply Markov's inequality to practical problems

Tchebyscheff's inequality is a powerful tool that we use in statistical analysis.
In this inequality, we remove the restriction that the random variable has to be non negative.
As a result, we now need to know additional information about the variable that is, (finite) expected value and (finite) variance.

In contrast to Markov's inequality, Tchebysheff allows you to estimate the deviation of the random variable from its mean. A common use of these inequalities is that it estimates the probability of the deviation from its mean in terms of its standard deviation.
Similar to Markov inequality, we can state **two** variants of Tchebysheff. Let us take a look at the simplest version.

2. Statement of the Inequalities Upper and Lower Bounds

Tchebysheff's inequality states as follows,
Let X be a Random variable for which Expected value of X and Variance of X exists. Then for any positive number k ,
Probability of modulus of (X minus mu) greater than or equal to k sigma is less than or equal to one by k square.
OR
Probability of modulus of (X minus mu) less than or equal to k sigma is greater than or equal to one minus 1 by k square.

By substituting k sigma with C, we get sigma is equal to C by k. Then the above inequality can be written as Probability of modulus of (X minus mu greater than or equal to C) is less than or equal to sigma square by C square.
OR
Probability of modulus of (X minus mu less than or equal to C) is greater than or equal to one minus sigma square by C square.

In Probability Theory Tchebysheff's inequality guarantees that in any probability distribution, "nearly all" values are close to the mean.

The precise statement being that no more than one by k square of the distribution's values can be more than K standard deviations away from the mean. Tchebysheff's Inequality, allows us to extend this idea to any distribution, even if that distribution isn't normal.

Upper Lower Bound:
The inequality probability of modulus of (X minus mu) greater than or equal to k sigma is less than or equal to one by k square provides an upper bound for the probability of the difference between a variable and its mean to be more than a constant k sigma.
Therefore one by k square is the upper bound.

Lower bound:
The inequality probability of modulus of (X minus mu) less than or equal to k sigma is greater than or equal to one minus one by k square, provides a lower bound for the probability of the difference between a variable and its mean to be less than a constant k sigma.
Therefore one minus one by k square is the lower bound.

Markov's Inequality
This is the simplest concentration inequality. Markov's inequality gives an upper bound for the probability that a non-negative function of a random variable is greater than or equal to some positive constant.

Let X be a non- negative random variable with finite expectation E(x) .Then for any epsilon greater than zero, probability of x greater than or equal to epsilon, is less than or equal to expected value of x by epsilon.

Markov's inequality takes mu is equal to expected value of x into account and provides an upper bound on probability of x greater than or equal to epsilon that depends on the values of epsilon and mu.

3. Illustrative Examples 1 & 2

Problem one:
Suppose we randomly select a journal article from a source with an average of one thousand words per article, with a standard deviation of two hundred words. Can we infer that the probability that it has between six hundred and one thousand four hundred words must be more than seventy five percent?

Solution:
Here mu is equal to one thousand and sigma is equal to two hundred and k is equal to two.
By Tchebyscheff's inequality, probability of modulus of (X minus mu) less than or equal to k sigma is greater than or equal to one minus one by k square.
This implies probability of, [minus k sigma less than or equal to (X minus mu) less than or equal to k sigma] is greater than or equal to one minus one by k square.
This implies probability of, [minus k into two hundred less than or equal to (X minus one thousand) less than or equal to k into two hundred] is greater than or equal to one minus one by k square.
This implies probability of, [one thousand minus k into two hundred less than or equal to X less than or equal to one thousand plus k into two hundred] is greater than or equal to one minus one by k square.

To find probability of, [six hundred less than or equal to X less than or equal to one thousand two hundred], we need to substitute k with two. On doing this, we get,
probability of [one thousand minus two into two hundred less than or equal to X less than or equal to one thousand plus two into two hundred] is greater than or equal to one minus one by four is equal to zero point seven five.
Which implies probability of [six hundred less than or equal to X less than or equal to one thousand four hundred] is greater than or equal to zero point seven five.
Which implies Probability of modulus of (X minus one thousand) less than or equal to two into two hundred is greater than or equal to zero point seven five.

Hence, by Tchebysheff's inequality, there is less than one by k square which is equal to one by four chance to be outside that range.
Or we can then infer that the probability that the article has between six hundred and one thousand and four hundred words (i.e. within k is equal to two Standard deviations of the mean) must be more than seventy five percent.

Problem 2:
Consider a coin that comes up with head with probability of zero point two. Let us toss it n times. Use Markov's inequality to bound the probability by at least eighty percent of heads.

Solution:
Let X be the random variable indicating the number of heads we got in n tosses. Clearly, X is non negative.
Using linearity of expectation, we know that Expected value of x is equal to zero point two into n.
We have to bound the probability by at least eighty percent which is equal to zero point eight.

Then we have to find Probability of X greater than or equal to zero point eight ‘n’. Using Markov’s inequality , we have,
Probability of (X greater than or equal to epsilon) is less than or equal to Expected value of X by epsilon.
Let epsilon be equal to zero point eight ‘n’. Then from Markov’s inequality we get
Probability of (X greater than or equal to zero point eight n) is less than or equal to zero point two ‘n’ by zero point eight ‘n’ which is equal to zero point two five.
Hence probability to get at least eighty percent of the heads is at most zero point two five.

4. Examples 3 & 4

Problem 3:
Two unbiased dice are thrown. If X is the sum of the numbers showing up, then prove that Probability of modulus of (X minus seven) greater than or equal to three is less than thirty five by fifty four. Compare this with the actual probability.

Solution

Figure 1

x	2	3	4	5	6	7	8	9	10
P(x)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36
x	11	12							
P(x)	2/36	1/36							

Let X denote the sum of the numbers showing up then, X takes values two , three, four, five, six, seven, eight, nine, ten, eleven and twelve.
The respective probabilities are, one by thirty six, two by thirty six, three by thirty six, four by thirty six, and five by thirty six, six by thirty six and so on for X is equal to twelve, and the probability is one by thirty six.

Then Expected value of X is equal to summation x into p of x which is equal to one by thirty six into [two plus six plus twelve plus twenty plus thirty plus forty two plus forty plus thirty six plus thirty plus twenty two plus twelve], which is equal to seven. Expected value of X square is equal to summation x square into probability of x which is equal to fifty four point eight three three.
Variance of X is equal to Expected value of x square minus expected value of x whole square which is equal to fifty four point eight three three minus seven square which is equal to five point eight three three.

Applying Tchebycheff's inequality Probability of modulus of (X minus mu greater than or equal to C) is less than sigma square by C square. This implies Probability of modulus of (X minus seven greater than or equal to C) is less than five point eight three three by C square. Here is equal to three. Probability of modulus of (X minus seven greater than or equal to three) is less than five point eight three three by nine which is equal to thirty five by fifty four.

To find the actual probability, probability of modulus of (X minus seven greater than or equal to three) is equal to Probability of (X minus seven) greater than three or minus of (X minus seven) is greater than or equal to three] which is equal to probability of (x greater than or equal to ten) plus probability of (x less than four). This is equal to probability of ten plus probability of eleven plus probability of twelve plus probability of two plus probability of three.
This is equal to three by thirty six plus two by thirty six plus one by thirty six plus one by thirty six plus two by thirty six.
This is equal to nine by thirty six which is equal to one by four. Hence the actual probability is

one by four.

Problem 4:
Seventy new jobs are opening up at an automobile manufacturing plant. One thousand applicants show up for seventy positions. To select the best seventy from among the applicants, the company gives a test to them. The mean and standard deviation of the scores turn out to be sixty and six respectively. Can a person who has a score of eighty four, count on getting one of the jobs?

Solution:
Given, mu is equal to sixty and sigma is equal to six.
Let X denote the score secured in the test.
Then by Tchebyscheff's inequality, probability of modulus of (X minus mu) greater than or equal to k sigma is less than or equal to one by k square. This implies probability of modulus of (X minus sixty) greater than or equal to k into six is less than or equal to one by k square. Let k be equal to four.

By substituting k equal to four, we get, probability of modulus of (X minus sixty) greater than or equal to twenty four is less than or equal to one by four square which is equal to one by sixteen. Probability of modulus of (X) greater than or equal to eighty four is less than or equal to one by four square which is equal to one by sixteen which is equal to zero point zero six two five. One thousand into zero point zero six two five is equal to sixty two point five.

Hence among one thousand applicants, around sixty three will have a chance of getting scores above eight four. Since there are seventy positions, a person who has a score of eighty four, has a chance of getting one of the jobs.

5. Examples 5, 6 & 7

Problem 5:
Let X be a Random variable taking values minus one, zero and one with probabilities one by eight, six by eight and one by eight respectively. Using Tschebyscheff's inequality, find the upper bound of the probability of modulus of x greater than or equal to one. Compare this with the actual probability.

Solution:
Here X takes values minus one, zero, and one with probabilities one by eight, six by eight and one by eight respectively. Then, Expected value of (X) is equal to summation x into probability of x which is equal to zero. Variance of x is equal to expected value of X square minus expected value of X whole square which is equal to one by four minus zero which is equal to one by four.
Therefore mu is equal to zero and sigma square is equal to one by four.

Probability of modulus of (x minus mu) greater than or equal to k sigma is less than or equal to one by k square. Substituting, mu is equal to zero and sigma is equal to one by two, we get, probability of modulus of (x) greater than or equal to k by two is less than or equal to one by k square .

Substituting k by two is equal one, we get k is equal to two. Therefore one by k square is equal to four. Therefore probability of modulus of (x) greater than or equal to one is less than or equal to one by four. The upper bound is equal to one by four.

To find the actual probability modulus of x greater than or equal to one means x greater than or equal to one. But x takes the values of minus one, zero and one. Therefore modulus of x greater than or equal to one means, x is equal to one and x is equal to minus one. Therefore probability of modulus of x greater than or equal to one is equal to probability of x is equal to one plus probability of x is equal to minus one which is equal to one by eight plus one by eight equal to one by four.
The upper bound is one by four and the actual probability is also one by four.

Problem 6:
X is a random variable with mean eight and variance four find a lower bound to probability of modulus of x minus eight less than four.

Solution:
By Tchebycheff's inequality
Probability of modulus of (X minus mu) less than or equal to k sigma is greater than or equal to one minus one by k square.
Given mu is equal to eight and sigma square is equal to four. Therefore probability of modulus of (x minus eight) less than or equal to k into two is greater than or equal to one minus one by k square.

Substituting two k is equal to four, we get k is equal to two. Then one minus one by k square is equal to one minus one by four which is equal to three by four

The lower bound of the probability is equal to one minus one by k square which is equal to three by four which is equal to zero point seven five.

Problem 7:

If X is a random variable with mean is equal to three and variance is equal to two. Find‘t’ such that probability of modulus of (x minus three) less than ‘t’ is greater than or equal to zero point nine nine.

Solution:

Given mu is equal to three and sigma square is equal to two. Therefore by Tchebycheff’s inequality, probability of modulus of (x minus mu) less than or equal to k into sigma is greater than or equal to one minus one by k square. Substituting probability of modulus of (x minus three) less than or equal to k into root two is greater than or equal to one minus one by k square. Call this as equation one.

Substitute one minus one by k square with zero point nine nine. Then one by k square is equal to one minus zero point nine nine which is equal to zero point zero one, which implies k is equal to ten. Substituting in equation one, we get, probability of modulus of (x minus three) less than or equal to ten into root two is greater than or equal to zero point nine nine. But given that, probability of modulus of (x minus three) less than‘t’ is greater than or equal to zero point nine nine. Therefore t is equal to ten into root two which is equal to fourteen point one four.

Here’s a summary of our learning in this session where we have :

- Explained the concept of Tchebyscheff’s and Markov’s inequalities
- Described the procedure to apply Tchebyscheff’s and Markov’s inequalities to practical problems
- Demonstrated the method to get the lower and upper bounds