M 362K Synopses for 2/17

Xiaohui Chen
EID: xc2388

February 16, 2015

The variance of a set of data measures the spread of a data set. Is is the difference between the second order expectation and the square of the first order expectation. For mean, $E[a \cdot X + b] = a \cdot E[X] + b$. However, $Var[a \cdot X + b] = a^2 \cdot Var[X]$. The transformation of variance is not affected by the translation constant b. The square root of variance is called the standard deviation. It represents how large the data are spread. We can use Chebychev's theorem to calculate the probability that a selected data is within some k standard deviation, where k is a constant larger than 1.

In order to compare two different data values, each of which is from distinct data sets, we can calculate the z-score, which is $z = \frac{X - \mu}{\sigma}$. By comparing the z-score, we can correctly compare two data values from two different data sets effectively.

The expectation and variance can also apply to random variables with conditional probabilities. The calculation method is the same as the ones with unconditional probabilities.