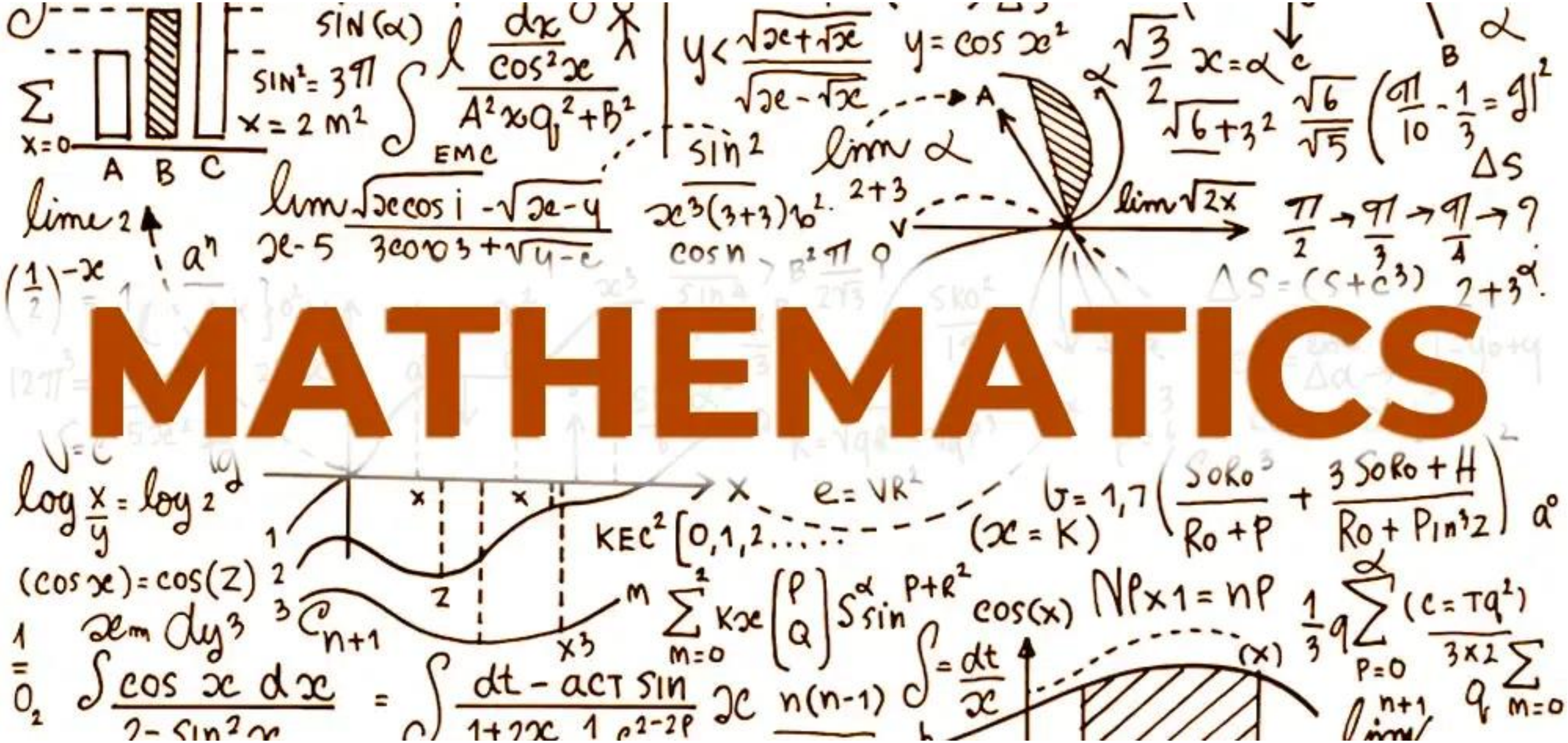

MATHEMATICS

Last lecture reminder

We learned about:

- Matrix basic operations - addition, subtraction, multiplication
- Introduction to scalars
- Introduction to vectors
- Vectors basic operations - addition, subtraction, multiplication
- Identity Matrix
- The transpose operator

The Inverse Matrix

Inverse → The inverse of a matrix is formed by creating a new matrix that, when multiplied with the original matrix, gives the identity matrix. The identity matrix is a square matrix in which all the elements of the principal (main or leading) diagonal are ones and all other elements are zeros. Note that not all matrices have inverses, and such matrices are called singular or degenerate.

The inverse matrix is usually denoted as **A^-1**, where A is the original matrix.

Determinant → In linear algebra, the determinant is a special number that can be calculated from a square matrix. The determinant helps us find the inverse of a matrix, tells us things about the matrix that are useful in systems of linear equations, calculus and more.

The Inverse Matrix

For each square matrix the determinant is calculated differently, so for example for 2X2 matrix

The determinant is calculated as A*D - B*C while A, B, C, D are the matrix values:

Matrix A (2X2)

[A, B]

[C, D]

Determinant of a matrix is often symbolized as **det(A)** or $|\mathbf{A}|$.

How the determinant help to find the inverse matrix?

A matrix has an inverse if, and only if, it is a square matrix (it has the same number of rows as columns) and its determinant is not equal to zero.

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})}\,\mathrm{adj}(\mathbf{A}).$$

adj(A) is the adjoint matrix
det(A) is the determinant of the original matrix

The Inverse Matrix

So we can also say that A matrix has an inverse if, and only if, it is a square matrix (it has the same number of rows as columns) and its determinant is not equal to zero.

Note: Every square matrix has a determinant but not every square matrix has a determinant that is not equal to 0. Because in order to find the inverse matrix we need to divided by the determinant value, matrix with determinants that are equal to 0 can’t have a inverse matrix.

For example → Let’s find the inverse matrix of the following 2X2 matrix

Matrix A (2X2):

[4, 7]

[2, 6]

det(A) = 4*6 - 7*2 = 24 - 14 = 10

The Inverse Matrix

We can see that we got a determinant that is not equal to 0 so there is an inverse matrix.

Now let's apply the formula to find this inverse matrix:

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \operatorname{adj}(\mathbf{A}).$$

So we calculated the det(A) and now we need to also calculated the adj(A).

Adjugate → The adjugate (or adjoint) of a matrix is the transpose of the cofactor matrix C of A. For a 2x2 matrix, adjugate matrix can be easily computed by swapping the diagonal elements and changing the signs of the off-diagonal elements.

So the adj(A) formula is the following:

adj(A) = [d, -b]

[-c, a]

We swapped the elements in the diagonal of the matrix (d, a) and we change the signs of the off-diagonal elements (b , c)

The Inverse Matrix

Back to our example, the adj(A) of the Matrix A (2X2) is → adj(A)

[4, 7]

[6 , -7]

[2, 6]

[-2, 4]

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \operatorname{adj}(\mathbf{A}).$$

now we have the determinant of t e matrix so we can finally calculate the inverse matrix:

A^-1 = 1/det(A) * [6 -7] = 1/10 * [6 -7] = [0.6 -0.7]

[-2 4] [-2 4] [-0.2 0.4]

We got that A^-1 = [0.6 -0.7]

[-0.2 0.4]

The Inverse Matrix

We can simply check our answer by performing A*A^-1 and see if we are getting the 2X2 identity matrix

A*A^-1 = [4 7] * [0.6 -0.7]   =   [(4*0.6 + 7*(-0.2))   (4*(-0.7) + 7*0.4)] =   [(2.4 - 1.4) (-2.8 + 2.8)]

[2 6]   [-0.2 0.4]   [(2*0.6 + 6*(-0.2))   (2*(-0.7) + 6*0.4)]   [(1.2 - 1.2) (-1.4 + 2.4)]

= [1 0]

[0 1]

Why the matrix inverse is important?

- **Solving Systems of Equations** → finding the inverse of the matrix can greatly simplify the task of finding the values of the variables of a linear equations represented by matrix.
- **Linear Transformations** → In the field of graphics and geometric transformations, inverse matrices are used to reverse transformations. If you have transformed a point in space using a transformation matrix, the inverse of the matrix can be used to transform the point back to its

Solving Linear Equations Using Inverse Matrix

The inverse of a matrix can be used to solve systems of linear equations represented in matrix form. The idea is that we can multiply both sides of the equation by the inverse of the matrix and "cancel out" the matrix on one side (because A * A-1 = I), leaving us with just the variables.

For example →

Let’s take our previsue matrix example:

Matrix A (2X2):

[4, 7 | 2]

[2, 6 | 1]

We learned that we can represent this as the following linear equation system:

$4x + 7y = 2$

$2x + 6y = 1$

Solving Linear Equations Using Inverse Matrix

We can also represent the same linear equation matrix as the following:

Matrix notation (AX = B):

[4, 7] [x] = [2]

[2, 6] [y]   [1]

Now let's multiply each side of the equation with A-1

We will get:

A * A-1* X = A-1 * B ⇒ I * X = A-1 * B ⇒ X = A-1* B

What we can see from this action is that finding the inverse matrix and multiply it with the result vector (B)

will give us the values of the unknown vector (X)

Solving Linear Equations Using Inverse Matrix

So now let’s solve the equation and find the unknown vector (X)

We already found the inverse matrix value for the original matrix:

Matrix A (2X2) → Inverse Matrix A-1 (2X2)

[4, 7] [0.6 -0.7]

[2, 6] [-0.2 0.4]

So now we need to calculate A-1 * B

[X] = [0.6 -0.7] * [2] ⇒ [x] = [0.6 * 2 + (-0.7 * 1))] ⇒ [0.5]

[Y] [-0.2 0.4] [1] [Y] [((-0.2 * 2) + 0.4 * 1)] [0]

So the solution for the linear equations is → x = 0.5 & y = 0

We can see that those numbers really solve the original linear equations so we got a right solution.

ECOM SCHOOL

המכללה למקצועות הדיגיטל וההייטק

Class Exercise - Solving Linear Equations

Instructions:

Solve the following linear equation by finding the inverse matrix and find the unknown vector values

- $2x + 5y = 12$
- $3x + 7y = 17$

Class Exercise Solution - Solving Linear Equations

Mathematical Concepts

Equation degree → The degree of an equation is the highest power of the variable in the equation.

Until now we only saw linear equations which only has variables to the power of 1, so the degree of linear

equation is 1. For example → $y = 2x + 3$ (first degree equation)

A **quadratic equation** contains a variable squared (to the 2nd power). So, the degree of a quadratic equation is 2. For example → $y = 2x^2 - x + 1$

A **cubic equation** contains a variable to the power of 3. So, the degree of a cubic equation is 3.

For example → $y = 2x^3 + 4x^2 - x + 1$

Note: We can go up in the degree of the equation as much as we like but solving higher degree equations can become much more complex as the degree gets higher. In addition the number of optional solutions get higher as well.

Mathematical Concepts

For example →

Let's take this simple linear equation: $3x + 2 = y$

The solution for $y = 8$ is $x = 2$, there are no more possible solutions.

Now let's take second degree equation: $x^2 + 2x + 1 = y$

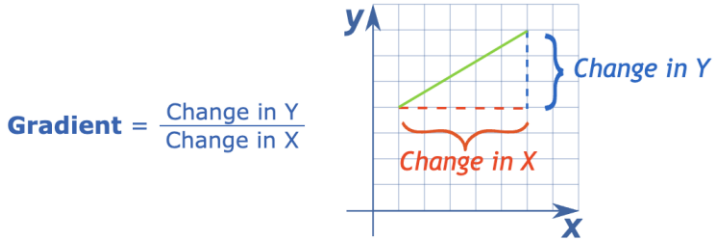
The solution for $y = 9$ is $x = 2$ and also $x = -4$

So for the second degree equation we found 2 different solutions that both of them solving the equation.

Gradient → The term "gradient" in mathematics refers to a measure of how a function changes as you move from one point to another. It's a concept used in calculus, particularly in differentiation, and is refer with the "slope" in simple algebra.

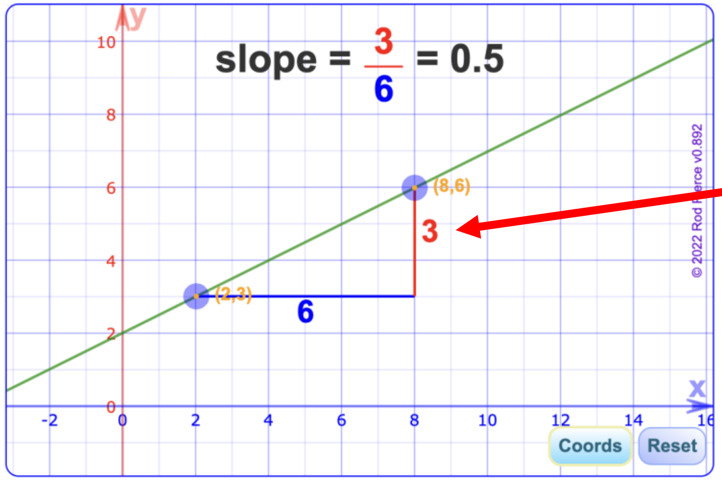
Mathematical Concepts

If we want to calculate the "slope" of a linear line between 2 points we need to calculate the change between the y points divided by the change between the x points:



This change can also been called **Delta** or **the symbol Δ (Delta)**

For example →



The gradient between the 2 points which mean the slope of the linear line is 0.5

Statistical Distribution

Distribution → In statistics, a distribution refers to the arrangement or pattern in which data values are spread across different ranges or categories in a dataset. It provides an overview of how the individual data points are distributed along a number line or within intervals.

There are two main types of statistical distributions:

Discrete Distribution → This is when the variable can take on only a finite number of values. Examples include rolling a die (results can only be between 1 and 6) or the number of calls at a call center (integer values only).

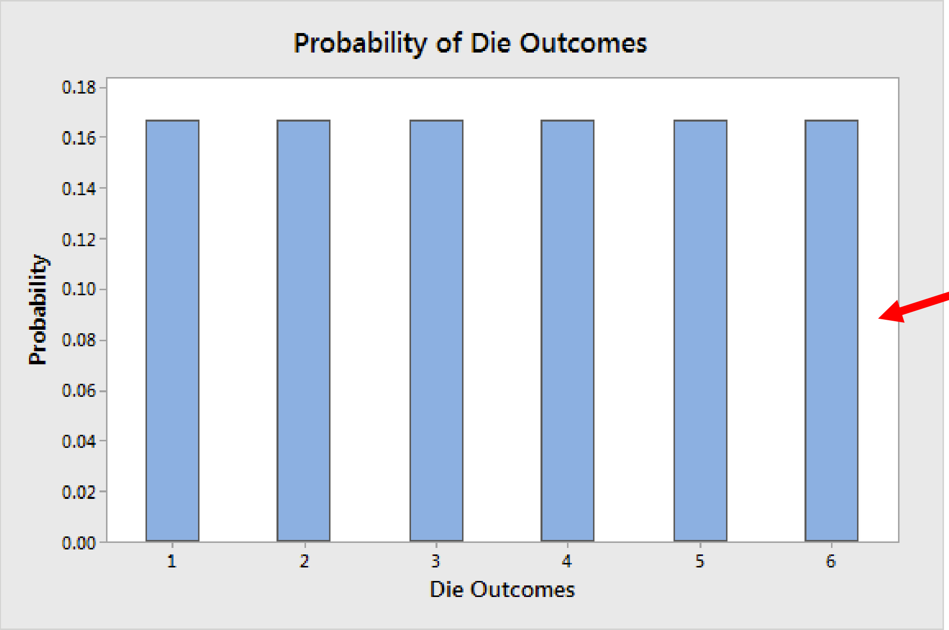
Continuous Distribution → This is when the variable can take on infinite values or any value in a range. Examples include weights of individuals or the time taken to run a mile (these can be any real number).

Famous District Distributions

Uniform distribution → Each outcome has the same probability of occurrence.

For example → The randint() function in Python is a uniform distribution.

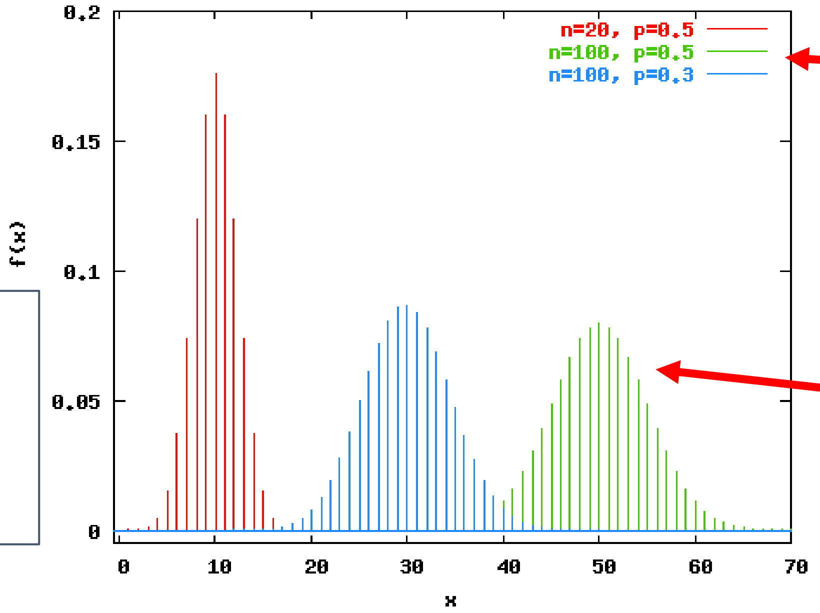
Another example is roll a fair dice which we can get each number with the same probability.



Each option has the same probability to get so the distribution between the different options is equal

Famous District Distributions

Binomial distribution → It models the number of successes in a fixed number of Bernoulli trials (a sequence of independent experiments with only two possible outcomes, called “success” and “failure”). **For example** → flipping a coin: A single coin flip can be considered a Bernoulli trial, which is a random experiment with exactly two possible outcomes, "success" and "failure". When we repeat the coin flip for a fixed number of times (say n times), and we are interested in the number of "successes" (number of "Heads", for example), then this follows a Binomial distribution.



P → The probability to get “success” in a single binomial try (should be represented as %, for example 50% chance to get success)
N → The number of binomial tries

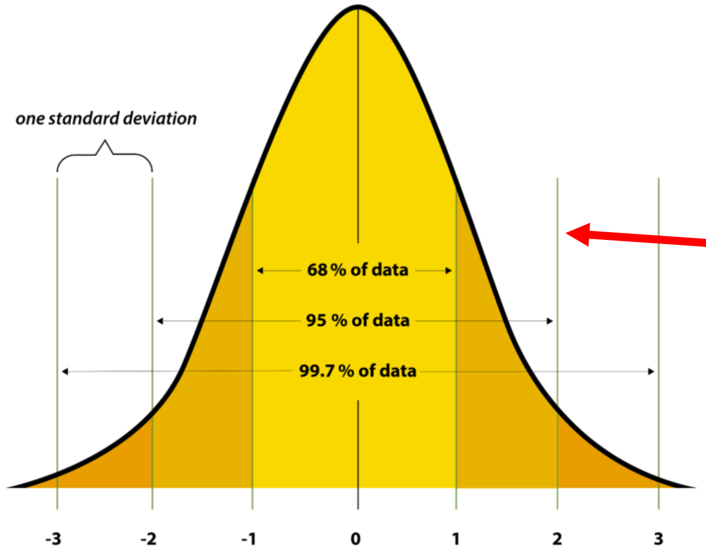
f(x) = The probability to get exactly x amount of success tries

x = Number of successes tries

We can see that the more trials we are doing the chance that we will get more successes is higher
In addition, the higher the probability to get success, the chance that we will get more successes is higher

Famous Continuous Distributions

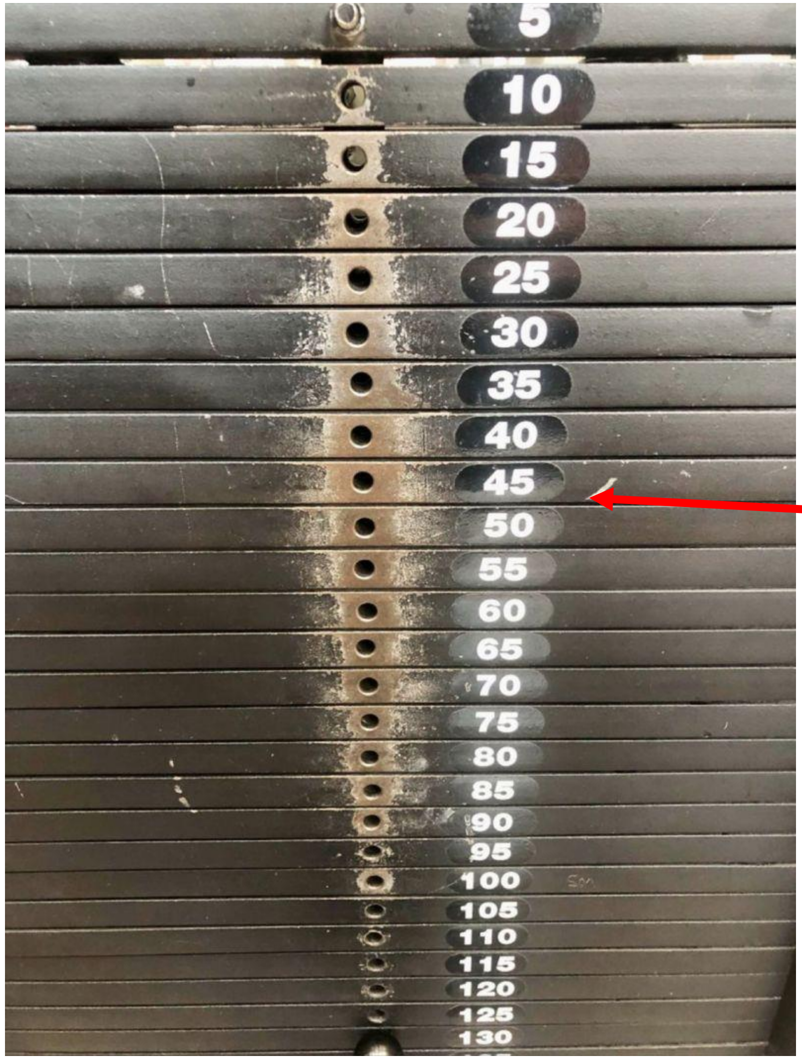
Normal distribution → The Normal distribution (also called Gaussian distribution) It's the classic, symmetrical, bell-shaped curve. Many natural phenomena follow normal distribution such as heights or IQ scores.



In normal distribution most of the people get the middle results. The further we move from the middle the lowest the chances to get those results.

ECOM SCHOOL
המכללה למקצועות הדיגיטל וההייטק

Famous Continuous Distributions



Most people in the gym are using the middle weight bars, the further we move from the middle weight bars the less the probability someone will use it.
This fit perfectly to normal distribution.

Statistical Concepts

Mean → The mean, often referred to as the average, is a fundamental concept in statistics. It represents the arithmetic sum of all the values in a data set divided by the count of values in the dataset.

Mean been represented as with μ (mu).

$\mu = (\Sigma X) / N$

Standard deviation → The standard deviation is a statistical measure that reflects the amount of variation or dispersion from the average (mean) in a data set. A low standard deviation indicates that the data points tend to be close to the mean, while a high standard deviation signifies that the data points are spread out over a wider range. SD been represented with σ (sigma).

$\sigma = \mathrm{sqrt}[(\Sigma(X - \mu)^2) / N]$

Σ symbolizes the sum

X represents each value in the dataset

μ denotes the mean (average) of the dataset

In order to calculate SD for a given dataset we first need to calculate the mean (μ), then for each point in the dataset we calculate it's distance from the mean value and square it. Finally, we calculate the average distance from the mean by sum every distance and divide by the number of values in the dataset.

Statistical Concepts

Variance → Variance is a statistical measurement that describes the spread of numbers in a dataset. More specifically, it measures how far each number in the set is from the mean (average) and thus from every other number in the set.

Variance can also be described as the squared standard deviation and been represented with σ^2 (sigma).

$\sigma^2 = (\Sigma(X - \mu)^2) / N$

For example → Let’s calculate the mean, SD and variance of this given dataset: [5, 7, 11, 15, 18]

Mean (μ) = (5 + 7 + 11 + 15 + 18) / 5 = 56 / 5 = 11.2

Now let’s calculate for each point the squared distance from the mean:

$(5-11.2)^2 = 38.44$

$(7-11.2)^2 = 17.64$

…

Statistical Concepts

Let’s do the same for all points and we will get that the variance is:

Variance (σ^2) = (38.44 + 17.64 + 0.04 + 14.44 + 46.24) / 5 = 116.8 / 5 = 23.36

Finally, let's calculate the standard deviation (σ):

Standard Deviation (σ) = sqrt(Variance) = sqrt(23.36) = 4.83 (rounded to two decimal places).

So, for the given dataset [5, 7, 11, 15, 18]:

Mean (μ) = 11.2

Variance (σ^2) = 23.36

Standard Deviation (σ) = 4.83

Class Exercise - Statistical Concepts

Instructions:

Calculate the mean, SD, and variance of the given dataset:

- [23, 45, 12, 37, 34, 55]

Class Exercise Solution - Statistical Concepts