Stats camp day 2

Frank Edwards

8/29/2019

Welcome back: math time!

Matrices

What is a matrix?

What is a matrix?

A matrix is a rectangular array of numbers, with dimensions expressed as rows $\times$ columns

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

$$B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$$

So A is a 3×3 matrix, B is a 3×2 matrix, and C is a 2×3 matrix.

Matrix notation

We can identify each element of a matrix with its column and row position, where x_{ij} refers to the value in the ith row and jth column of matrix X. Note that we use uppercase letters for a matrix, and lowercase letters for elements of a matrix.

$$\mathsf{X} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{bmatrix}$$

Vectors are one-dimensional arrays of values. Either an n-row column or an n-column row:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix}$$

OR

$$\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

Special matrices

A diagonal matrix has zero values except on the diagonal:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 9 \end{bmatrix}$$

An identity matrix is a special case of a diagonal matrix, where all values on the diagonal are equal to 1

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

These matrices are also symmetric, where all $x_{ij} = x_{ji}$. Symmetric matrices must be square.

Matrix operations

Matrix transpose

A transpose interchanges the rows and columns of a matrix, or rotates it.

The dimensions are switched, so an $n \times k$ matrix becomes a $k \times n$ matrix.

We denote a transpose with a T

$$\mathrm{B} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \mathrm{B}^{\mathsf{T}} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

Addition and subtraction

Two matrices (or vectors) can be added or subtracted only if they have identical dimensions. Then add or substract the correspoding elements of the matrices.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 1+5 & 2+6 \\ 3+7 & 4+8 \end{bmatrix}$$

Multplication by scalar

Matrices and vectors can be multiplied by constant values (called scalars).

$$\mathsf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} 4\mathsf{B} = \begin{bmatrix} 4 & 8 & 12 \\ 16 & 20 & 24 \end{bmatrix}$$

$$\mathsf{C} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} 3\mathsf{C} = \begin{bmatrix} 3 \\ 6 \\ 9 \\ 12 \end{bmatrix}$$

Matrix multiplication

Two matrices can be multiplied only if the numbers of columns in the first matrix is equal to the number of rows of the second matrix. The resulting matrix has a number of rows of the first matrix and a number of columns of the second matrix.

$$\mathsf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \mathsf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}$$

We find each element (ab_{ij}) by summing the cross product of the ith row of A and the jth column of B.

$$A \cdot B = \begin{bmatrix} a_{11} \cdot b_{11} + a_{12} \cdot b_{21} + a_{13} \cdot b_{31} & a_{11} \cdot b_{12} + a_{12} \cdot b_{22} + a_{13} \cdot b_{32} \\ a_{21} \cdot b_{11} + a_{22} \cdot b_{21} + a_{23} \cdot b_{31} & a_{21} \cdot b_{12} + a_{22} \cdot b_{22} + a_{23} \cdot b_{32} \end{bmatrix}$$

Exercises

Use the following matrix for these exercises:

$$\mathsf{A} = \begin{bmatrix} 2 & 7 & 7 \\ 8 & 3 & 1 \\ 0 & 4 & 6 \end{bmatrix}$$

1. Identify $A_{1,1}$
2. Identify $A_{2,3}$
3. Provide A^T
4. Compute $A + A$
5. Compute $3A$

Break

Probability

Set notation

A set is a collection of elements from a population

- Positive integers ≤ 5: $A = \{1, 2, 3, 4, 5\}$
- Pizza toppings: $B = \{\text{mushrooms}, \text{spinach}, \text{peppers}\}$

Set notation

A set is a collection of elements from a population

- Positive integers ≤ 5: $A = \{1, 2, 3, 4, 5\}$
- Pizza toppings: $B = \{\text{mushrooms}, \text{spinach}, \text{peppers}\}$

A set is *empty*, written as $\varnothing$, if it contains no elements. A set is *universal*, written as Ω, if it contains all elements in the population.

Sets: intersection

The intersection of two sets A and B is the set of all elements that are in both A and B. We write the intersection of A and B as $A \cap B$.

$$A = \{1, 2, 3, 4\}, B = \{2, 4, 6, 8\}, A \cap B = \{2, 4\}$$

Sets: intersection

The intersection of two sets A and B is the set of all elements that are in both A **and** B. We write the intersection of A and B as $A \cap B$.

$$A = \{1, 2, 3, 4\}, B = \{2, 4, 6, 8\}, A \cap B = \{2, 4\}$$

$$A = \{\text{Odd numbers}\}, B = \{\text{Even numbers}\}, A \cap B = \{\varnothing\}$$

Sets: union

The union of two sets A and B is the set of all elements that are in A **or** B.

We write the union of A and B as $A \cup B$

$$A = \{1, 2, 3, 4\}, B = \{2, 4, 6, 8\}, A \cup B = \{1, 2, 3, 4, 6, 8\}$$

Sets: union

The union of two sets A and B is the set of all elements that are in A or B.
We write the union of A and B as $A \cup B$

$$A = \{1, 2, 3, 4\}, B = \{2, 4, 6, 8\}, A \cup B = \{1, 2, 3, 4, 6, 8\}$$

$$A = \{\text{Odd numbers}\}, B = \{\text{Even numbers}\}, A \cup B = \{\text{All integers}\}$$

Sets: subset

If all elements of A are also in B, then A is a subset of B.

- $A = \{1, 4\}, B = \{1, 2, 3, 4\}$: A is a subset of B. B is not a subset of A.

Sets: mutually exclusive

If the intersection of A and B is $\varnothing$ (they have no elements in common), then A and B are mutually exclusive

- $A = \{1, 5\}, B = \{2, 3\}$
- $A = \{\text{Odd numbers}\}, B = \{\text{Even numbers}\}$

Sets: complement

The complement of a set A is the set of elements in the population that are not in A, denoted by A^c

- Population: Integers 1 - 10: $A = \{1, 7\}, A^c = \{2, 3, 4, 5, 6, 8, 9, 10\}$

Sample space

An experiment is an action or process of observation. It has only one outcome, but it is not known with certainty until carried out. Flipping a coin or rolling a dice is an experiment.

Sample space

An experiment is an action or process of observation. It has only one outcome, but it is not known with certainty until carried out. Flipping a coin or rolling a dice is an experiment.

The sample space is the set with all possible outcomes of the experiment. For a coin flip, the sample space $S = \{\text{Heads, Tails}\}$

Sample space

We conduct a simple experiment: flip a fair coin three times.

There are two possible outcomes for each coin.

With two possible outcomes on each flip, the total number of possible outcomes is $2 \times 2 \times 2 = 8$

The sample space is $S = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$

Sample space

We conduct a simple experiment: flip a fair coin three times.

There are two possible outcomes for each coin.

With two possible outcomes on each flip, the total number of possible outcomes is $2 \times 2 \times 2 = 8$

The sample space is $S = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$

Sample space

What is the sample space for rolling 2 standard six sided dice?

Sample space

What is the sample space for rolling 2 standard six sided dice?

$$S = \{1,1;1,2;1,3;1,4;1,5;1,6;2,1;2,2;2,3...\}$$

How many possible outcomes are there?

Sample space

What is the sample space for rolling 2 standard six sided dice?

$$S = \{1, 1; 1, 2; 1, 3; 1, 4; 1, 5; 1, 6; 2, 1; 2, 2; 2, 3 \ldots\}$$

How many possible outcomes are there?

6 possible outcomes for the first dice $\times$ 6 possible outcomes for the second dice $= 36$

Events

An event is a subset of the sample space.

An event is a subset of the sample space.

For our 3 coin flip experiment, the sample space is:

$$S = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$$

Events

An event is a subset of the sample space.

For our 3 coin flip experiment, the sample space is:

$S = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$

Events we could observe:

- Flipping two heads: HHT, HTH, THH
- Getting an odd number of tails: HHT, HTH, THH, TTT
- More than one heads: HHH, HHT, HTH, THH

Probability is the likelihood of observing an outcome. We compute probability by counting the number of ways a specific outcome could occur, then divide by the total number of possible outcomes. The sum of the probability of each outcome in the sample space must be equal to 1.

Probability is the likelihood of observing an outcome. We compute probability by counting the number of ways a specific outcome could occur, then divide by the total number of possible outcomes. The sum of the probability of each outcome in the sample space must be equal to 1.

For a six-sided dice, the sample space is $S = \{1, 2, 3, 4, 5, 6\}$, and each outcome has the same chance of occurring.

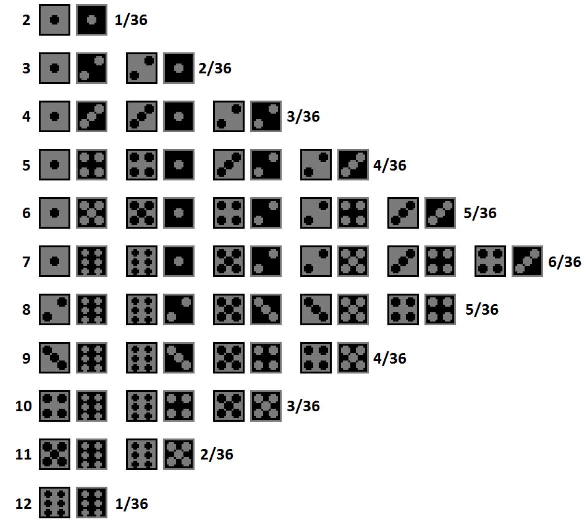
Probability

Probability is the likelihood of observing an outcome. We compute probability by counting the number of ways a specific outcome could occur, then divide by the total number of possible outcomes. The sum of the probability of each outcome in the sample space must be equal to 1.

For a six-sided dice, the sample space is $S = \{1, 2, 3, 4, 5, 6\}$, and each outcome has the same chance of occurring.

What is the probability of rolling a 3?

What is the probability of each number 1-12 when rolling two dice?



Probability of events

The probability of an event is the sum of the probability of each of the elements in the event.

Probability of events

The probability of an event is the sum of the probability of each of the elements in the event.

- What is the probability of a dice roll producing a value less than or equal to 5?

Probability of events

The probability of an event is the sum of the probability of each of the elements in the event.

- What is the probability of a dice roll producing a value less than or equal to 5?

$$A = \{1, 2, 3, 4, 5\}, P(A) = ?$$

Probability of events

The probability of an event is the sum of the probability of each of the elements in the event.

- What is the probability of a dice roll producing a value less than or equal to 5?

$$A = \{1, 2, 3, 4, 5\}, P(A) = ?$$

$$P(1) + P(2) + P(3) + P(4) + P(5) = 1/6 + 1/6 + 1/6 + 1/6 + 1/6 = 5/6$$

- What is the probability that a dice roll is even?

$$B = \{2, 4, 6\}, P(B) = P(2) + P(4) + P(6) = 1/6 + 1/6 + 1/6 = 3/6$$

If we know the probability of observing an element in set A, then we also know the probability of observing something in the set that is the complement of A, A^c.

$$P(A^c) = 1 - P(A)$$

Probability of the complement

If we know the probability of observing an element in set A, then we also know the probability of observing something in the set that is the complement of A, A^c.

$$P(A^c) = 1 - P(A)$$

The probability of observing an event is 1 minus the probability of not observing the event.

Conditional probability

Conditional probability is the probability of observing one event when we know that another event already occurred.

Conditional probability is the probability of observing one event when we know that another event already occurred.

If we know that a dice roll is even ($B = \{2, 4, 6\}$), what is the probability that it was a 2 ($A = \{2\}$)?

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Conditional probability

The outcome in both A *and* B is 2.

$$A \cap B = \{2\}$$

$$P(A \cap B) = P(2) = 1/6$$

The probability of rolling an even is $P(2) + P(4) + P(6) = 3/6$

The outcome in both A *and* B is 2.

$$A \cap B = \{2\}$$

$$P(A \cap B) = P(2) = 1/6$$

The probability of rolling an even is $P(2) + P(4) + P(6) = 3/6$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/6}{3/6} = 1/3$$

Independence

If knowing B does not give is any information about A, then A and B are independent.

Using conditional probability, if P and A are independent:

$$P(A|B) = P(A)$$

Bayes' Rule

For two events, if we know one conditional probability for the two events, and the overall probability of each event (the marginal probabilities), we can use Bayes' rule to compute the other conditional probability.

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

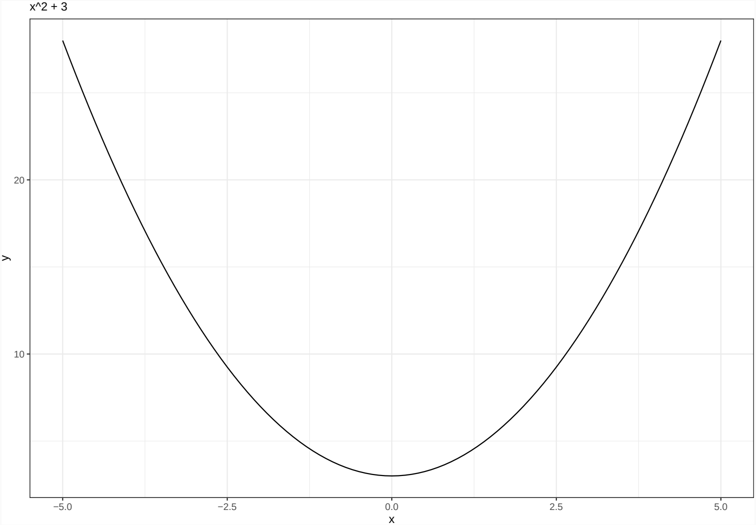
Exercises

1. What is the sample space for the result of two coin flips?
2. What is the probability of seeing one heads, one tails?
3. $A = \{1, 3, 4\}, B = \{2, 4, 6\}, A \cap B = ?$
4. $A = \{1, 3, 4\}, B = \{2, 4, 6\}, A \cup B = ?$
5. $P(A \cap B) = 0.4, P(B) = 0.2, P(A|B) = ?$

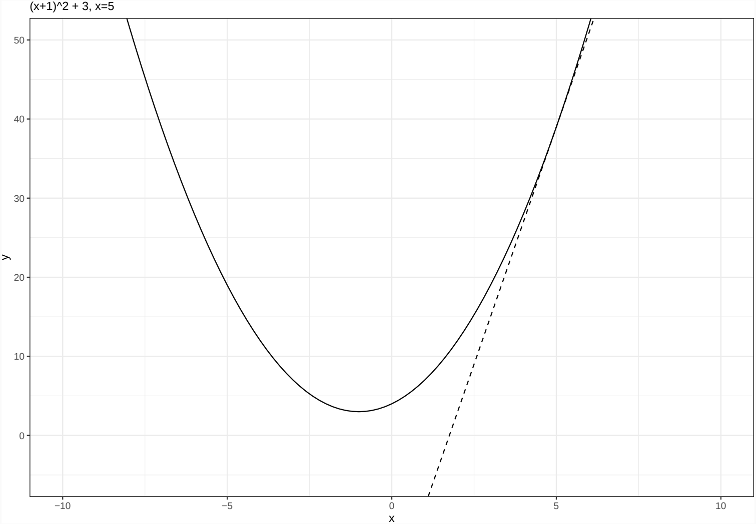
Break

Calculus

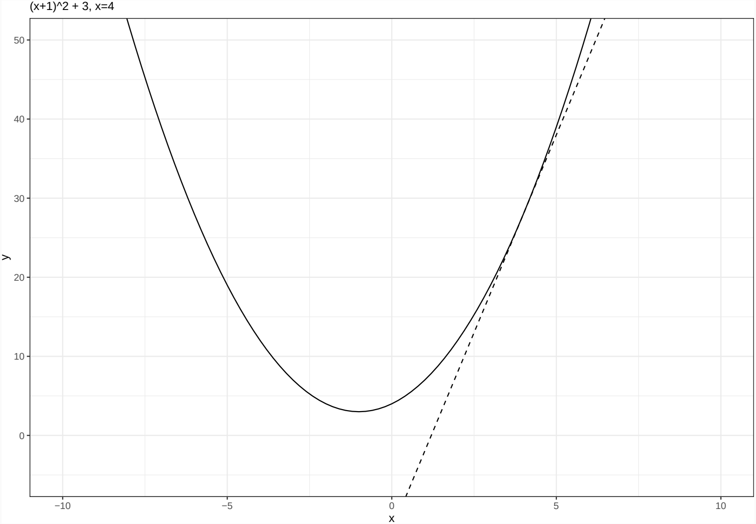
What is the derivative of a curve?



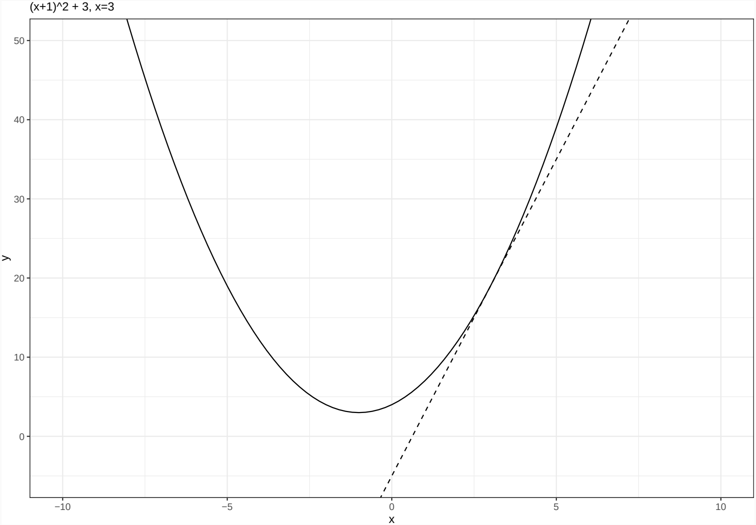
What is the derivative of a curve?



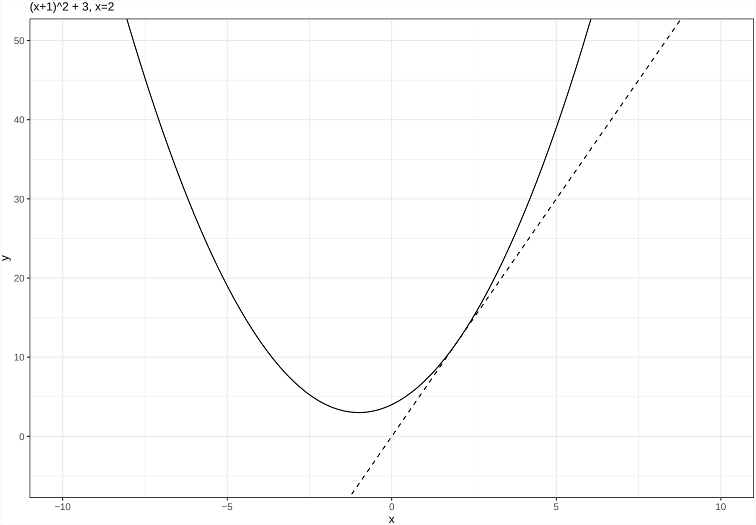
What is the derivative of a curve?



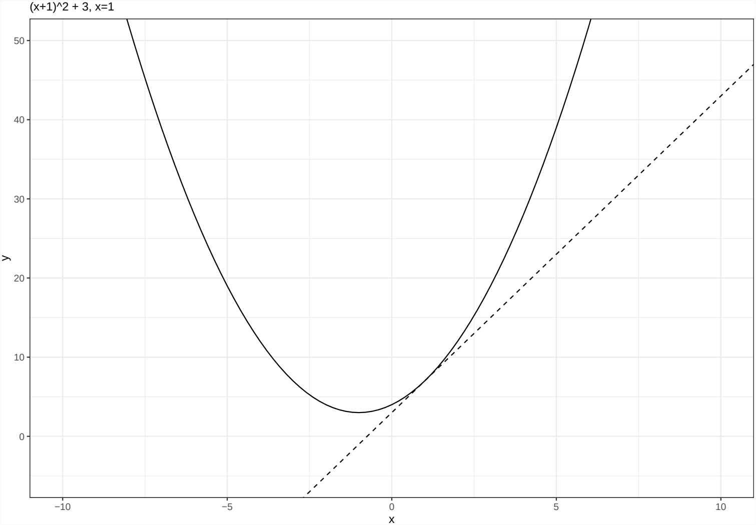
What is the derivative of a curve?



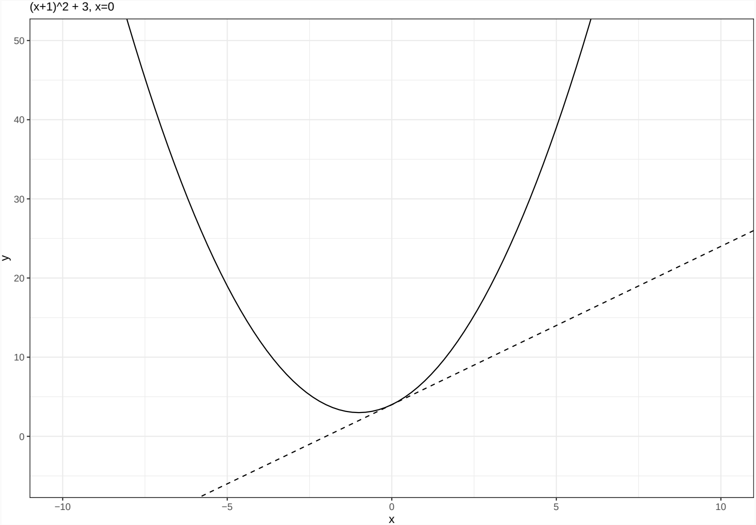
What is the derivative of a curve?



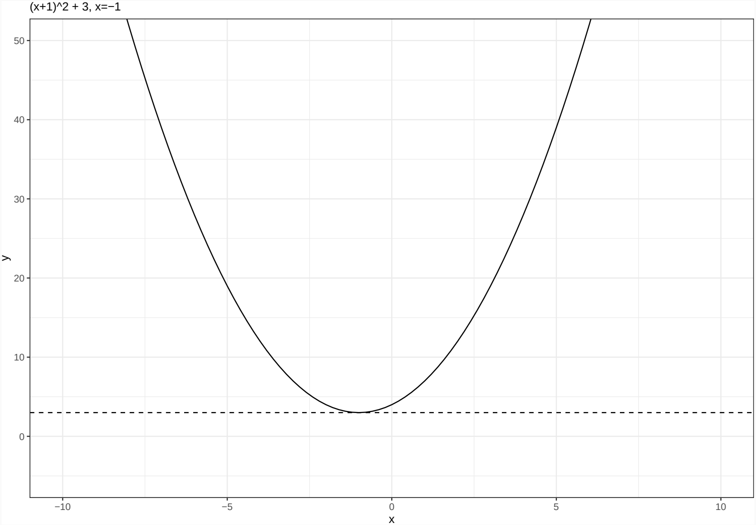
What is the derivative of a curve?



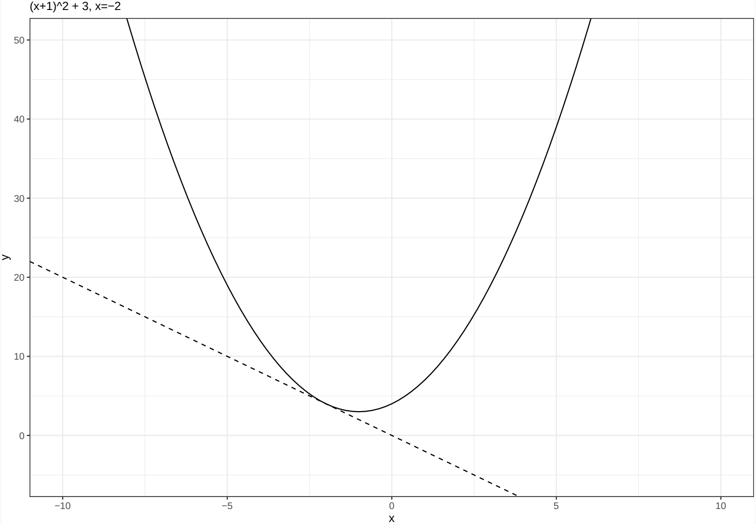
What is the derivative of a curve?



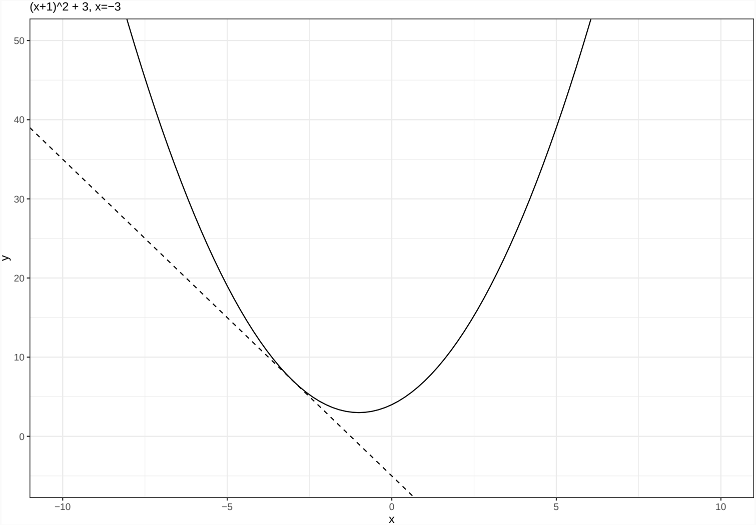
What is the derivative of a curve?



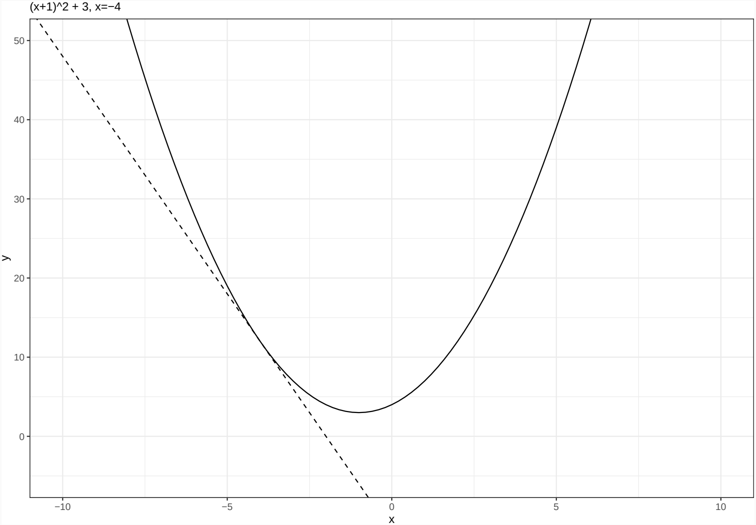
What is the derivative of a curve?



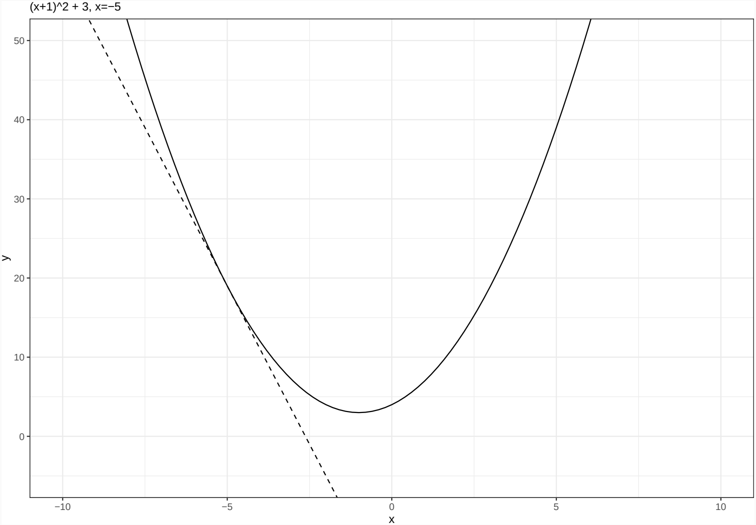
What is the derivative of a curve?



What is the derivative of a curve?



What is the derivative of a curve?



Finding the derivative

- The derivative of the function $f(x)$ is the instantaneous rate of change of the function at x.
- Equivalent to finding the slope of the function at a specific point

- The derivative of the function $f(x)$ is the instantaneous rate of change of the function at x.
- Equivalent to finding the slope of the function at a specific point

For two points b and a, we can find a slope as

$$\frac{f(b) - f(a)}{b - a}$$

Finding the derivative

If we want to know the *instantaneous* rate of change, then we should add a small value δ to x to see how much the value of $f(x)$ changes relative to $f(x+\delta)$.

$$\frac{f(x+\delta)-f(x)}{\delta}$$

The derivative $\frac{d}{dx}f(x)$ or $f'(x)$ is

$$lim_{\delta\to 0}\frac{f(x+\delta)-f(x)}{\delta}$$

Differentiation rules

Derivative of a constant:

$$f(x) = a; f'(x) = 0$$

Derivative of a power:

$$f(x) = ax^n; f'(x) = n \cdot a \cdot x^{n-1}$$

Derivative of an exponential:

$$f(x) = e^x; f'(x) = e^x$$

Examples of derivatives

- $$f(x) = x^5; f'(x) = 5x^4$$

- $$f(x) = 2x^3 + 3x + 9; f'(x) = 6x^2 + 3$$

- $$f(x) = 2e^x; f'(x) = 2e^x$$

Exercises

- $f(x) = x^3; f'(x) = ?$
- $f(x) = 2; f'(x) = ?$
- $f(x) = 3x; f'(x) = ?$
- $f(x) = 4x^2 + 2x + 1; f'(x) = ?$

Using the derivative to find critical values

We can use derivatives to find places where functions change from increasing to decreasing.

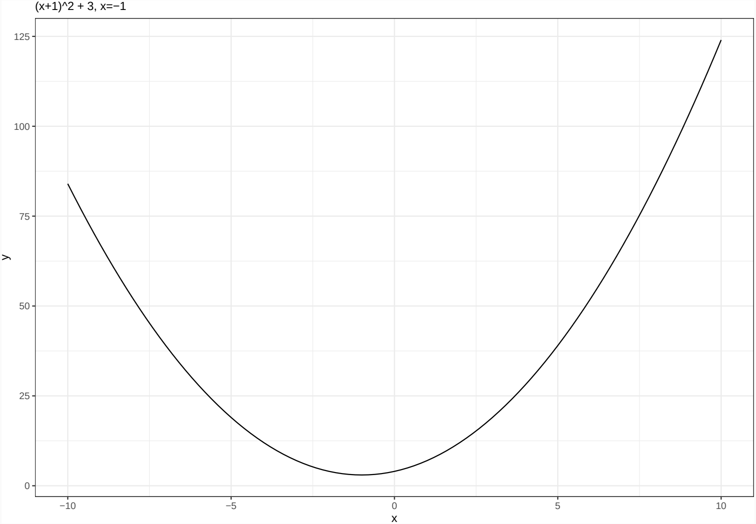
Using the derivative to find critical values

We can use derivatives to find places where functions change from increasing to decreasing.

- A maximum is a point where a function stops increasing and starts to decrease
- A minimum is a point where a function stops decreasing and starts to increase

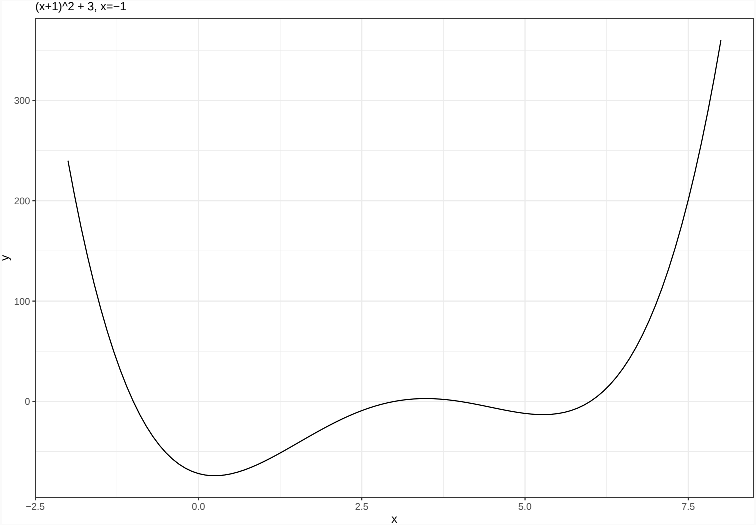
Critical points are defined as points where the derivative is zero.

Where is the critical point? Is it a maximum or minimum?



Where is the critical point? Is it a maximum or minimum?

Local maxima and minima



Integration

Distance, velocity, acceleration

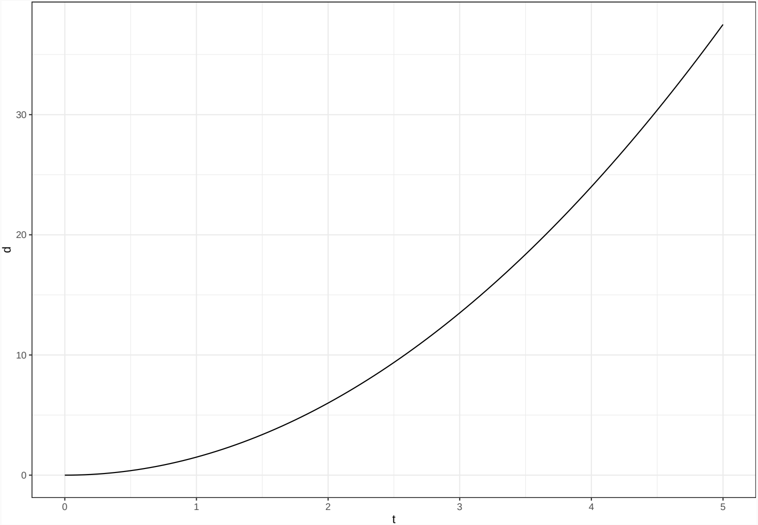
$d =$ distance, $v =$ velocity, $a =$ acceleration, $t =$ time

$$d(t) = \frac{a}{2}t^2$$

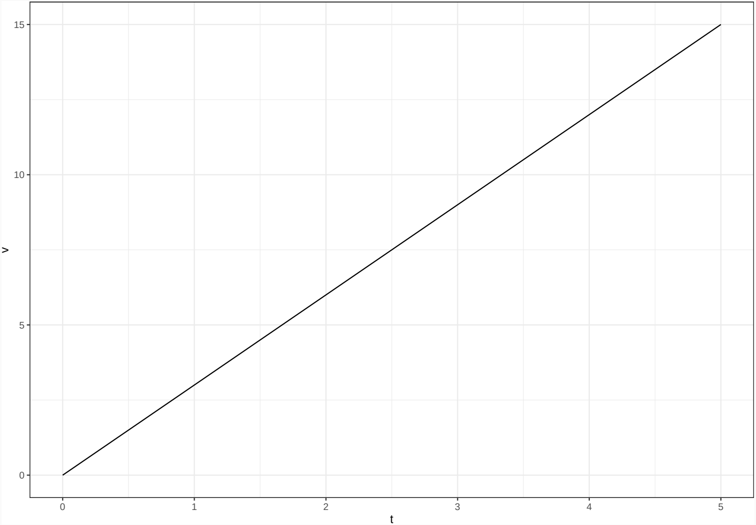
$$v(t) = d'(t) = 2 \cdot \frac{a}{2}t = at$$

$$a(t) = v'(t) = d''(t) = a$$

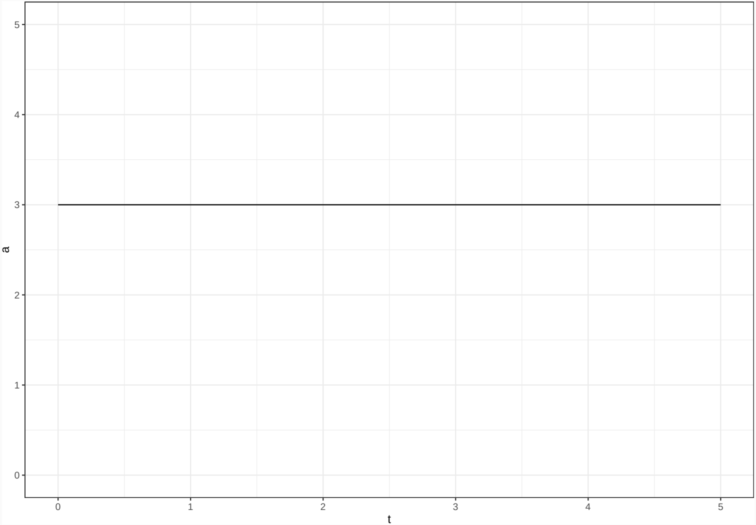
Distance as a function of time: a = 3



Velocity as a function of time: a = 3



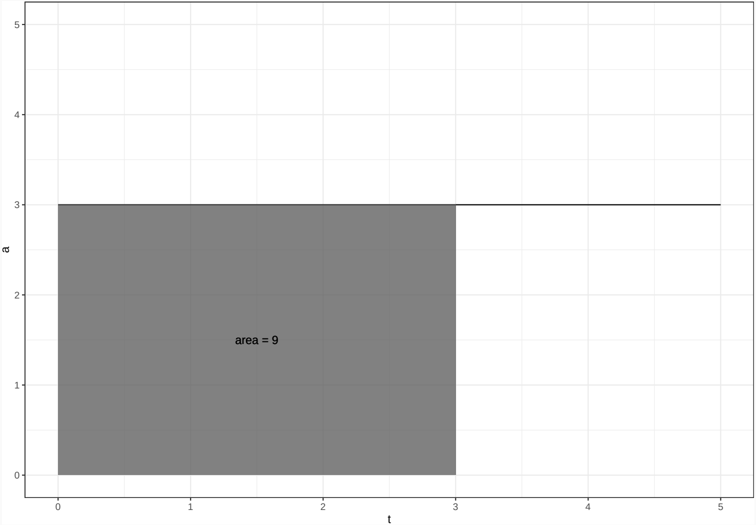
Acceleration as a function of time: $a = 3$



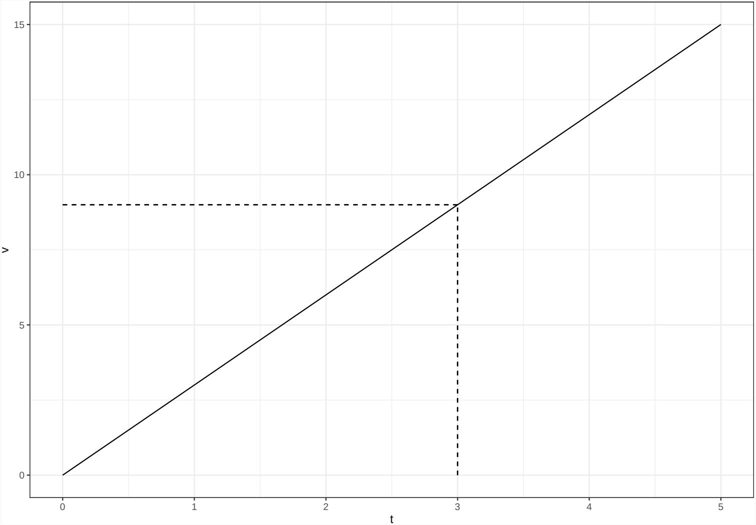
What is the velocity when t = 3

We know that acceleration is a constant of 3 here, so velocity at time 3 is simply $3 \cdot 3$. But we can also find the area under the acceleration curve from zero to 3 to find the velocity.

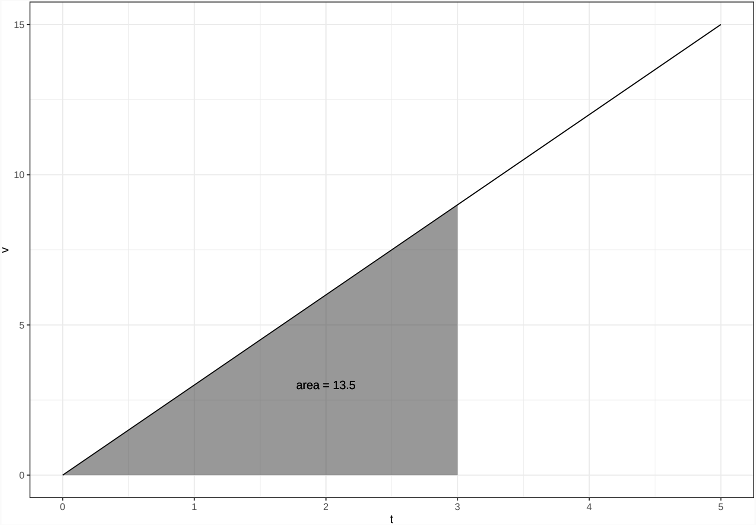
What is the velocity when t = 3



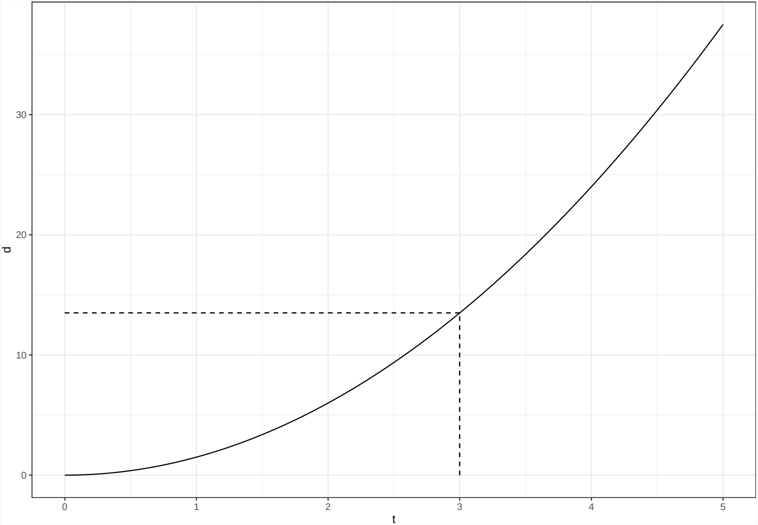
What is the velocity when t = 3?



What is the distance when t = 3?



What is the distance at time = 3?



We write the integral of a function from point a to point b as

$$\int_a^b f(x)dx$$

This provides us with the area under the curve $f(x)$ for values of x between a and b.

Mathematical notation

We write the integral of a function from point a to point b as

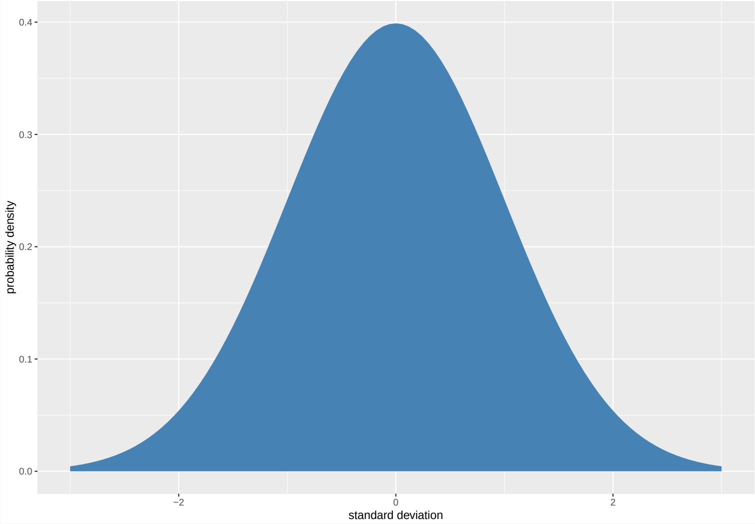
$$\int_a^b f(x)dx$$

This provides us with the area under the curve $f(x)$ for values of x between a and b.

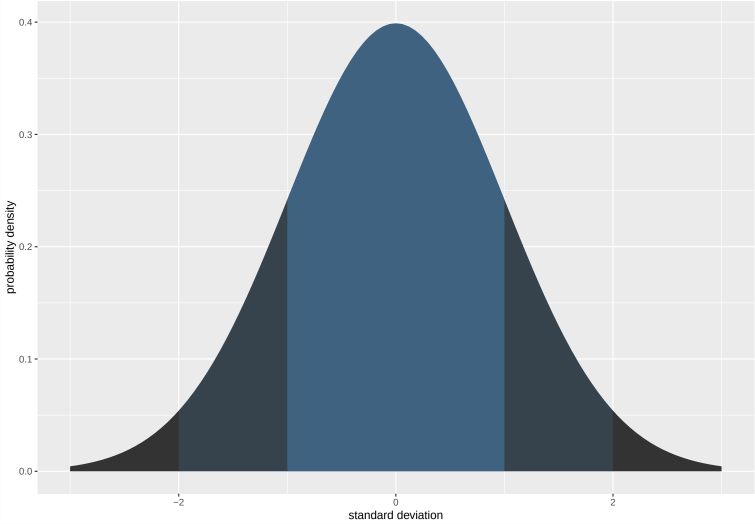
The integral is the inverse of the derivative.

$$\int f'(x)dx = f(x)$$

Why do we need this for statistics?



Why do we need this for statistics?



Questions?

Acknowledgments

These slides borrow heavily from Laina Mercer's math camp slides for University of Washington CSSS doctoral students. That course is available here: `https://www.csss.washington.edu/academics/math-camp/lectures`