# A Very Brief Review of Single & Multi-variate Calculus

DSI Week 0 Lecture - January 19, 2018

galvanize

# Learning Objectives

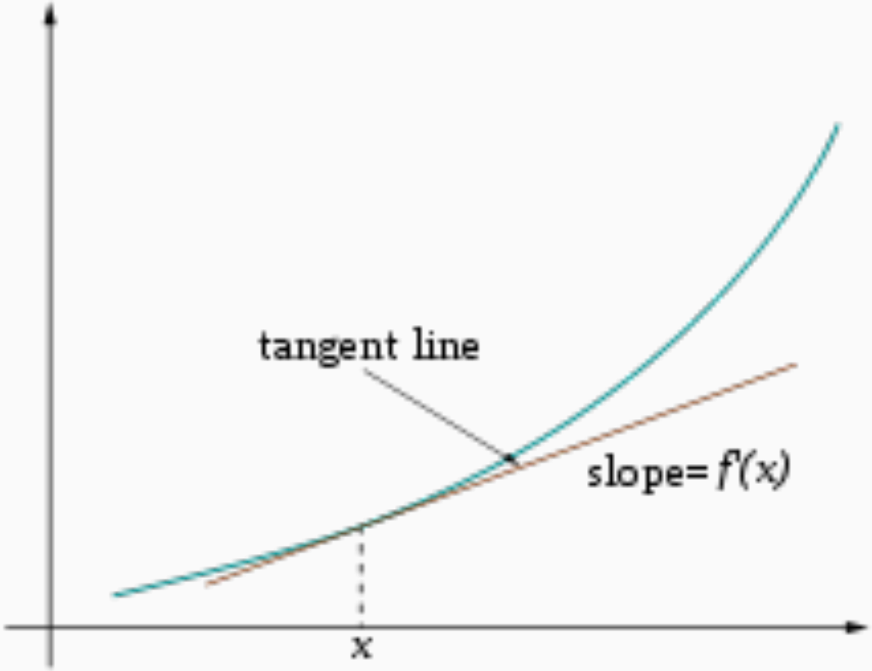
- Review basic calculus operations:
  - single variable derivatives
  - Integrals
  - Partial derivatives / multivariable derivatives
- Motivation : applications of the above

# Single Variable Derivatives

galvanize

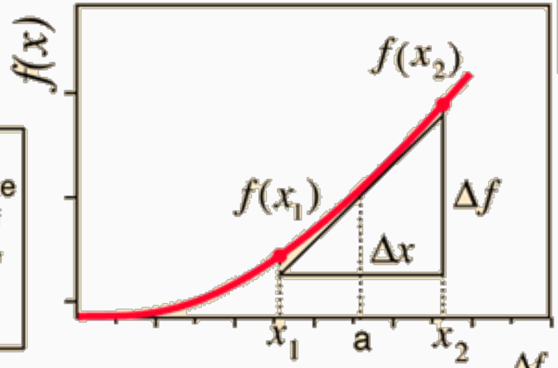
A derivative is an instantaneous representation of the slope of a function at point x

The slope of the line tangent to f(x) at point x



# Single Variable Derivatives

The derivative of f(x) with respect to x is the instantaneous rate of change (slope) of the function at any value of x.




The $\boldsymbol{d}$ in both numerator and denominator denotes the derivative.

$$\frac{df(x)}{dx} = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x}$$

$\boldsymbol{x}$ is the independent variable upon which the function $\boldsymbol{f(x)}$ depends.

The slope of the line tangent to the curve f(x) is given by $\frac{\Delta f}{\Delta x}$

That slope is also given approximately by $\frac{f(x_2) - f(x_1)}{x_2 - x_1} = \frac{\Delta f}{\Delta x}$ and the approximation would be better if $x_2$ and $x_1$ were closer together, i.e., if $\Delta x$ were smaller. The exact slope can be found by taking the limit as $\Delta x \to 0$ and that limit is called the derivative of the function f(x) with respect to x.

The derivative gives the instantaneous rate of change of the function.

$$\frac{df(x)}{dx} = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x}$$

The value of the derivative at x=a is equal to the slope fo the line tangent to the curve f(x) at the point (a, f(a)).

# Recall: Product Rule!

$$(f \cdot g)' = f' \cdot g + f \cdot g'$$

Whiteboard example: $f(x) = x \ln(x)$

now you try… $f(x) = x^2 e^{-x}$

# Recall: Chain Rule!

$$f(g(x))' = f'(g(x)) \cdot g'(x)$$

Whiteboard example: $$f(x) = \sqrt{\frac{1}{2\pi}} e^{-x^2}$$

now you try… $$f(x) = (\ln(x))^2$$

# Applications

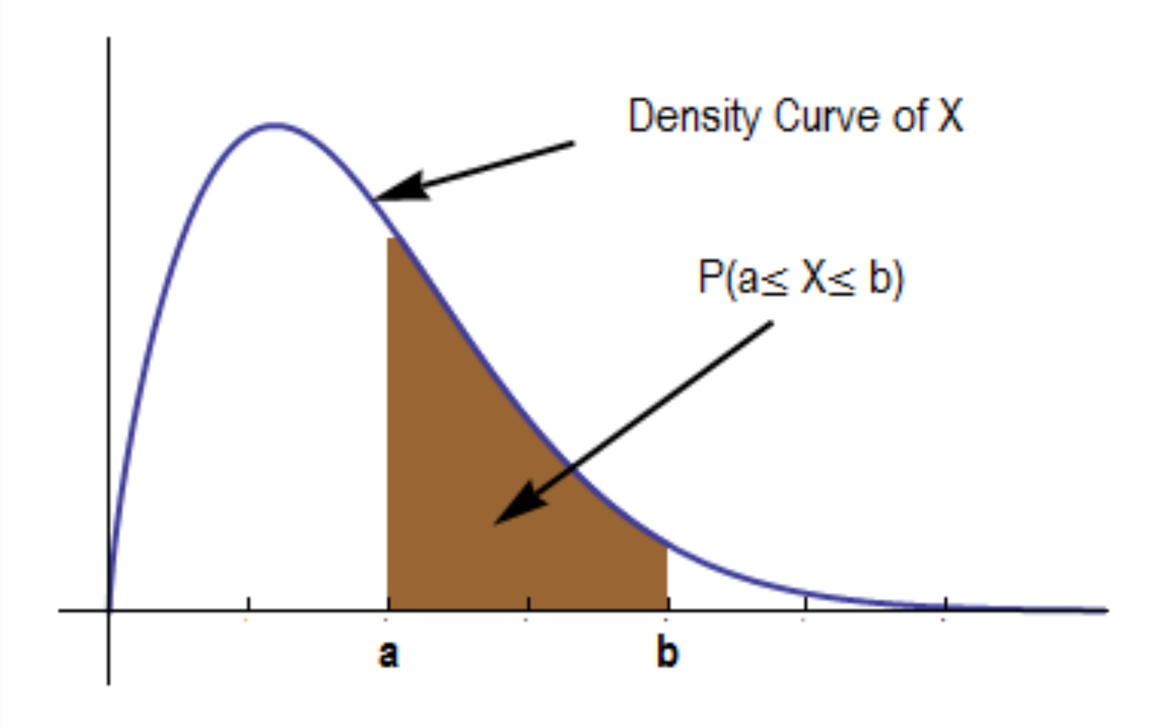


Maximizing and Minimizing things…

- Maximum Likelihood Estimates (MLE) for parametric estimators
- (mostly….we live in partial derivative world)

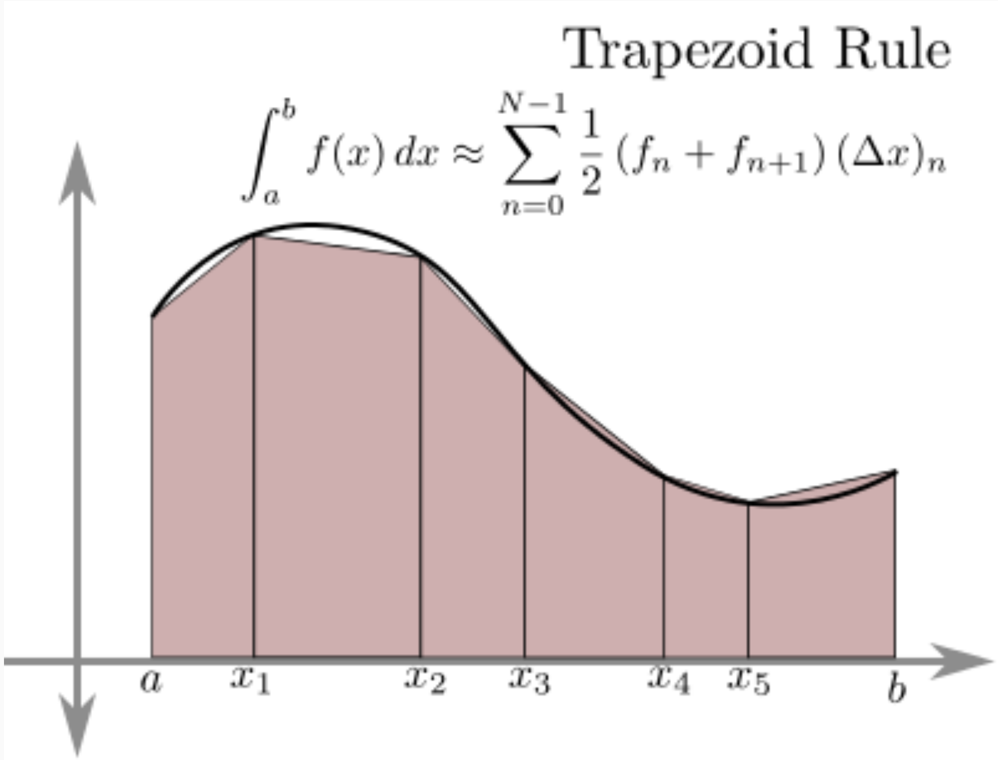
# Integration

AKA anti derivatives! We are trying to find the area under the curve



# Numerical Integration

For non-integrable functions (and really even some things that would just be a pain to integrate analytically….) can approximate area under the curve using rectangles or trapezoids



# Numerical Integration

It is unlikely you will need to implement this, but if you do, it would like something like this…

```
# pseudocode
# start area under the curve = 0
AUC = 0
n_partitions = (xmax - xmin)/deltax

for i in range(n_partitions):
    # figure out the min and max for this piece
    a = xmin + i*deltax
    b = xmin + (i+1)*deltax

    # do trapezoidal area
    trap_area = 0.5 * (func(a) + func(b)) * deltax

    AUC += trap_area
```

# Integration Applications

galvanize

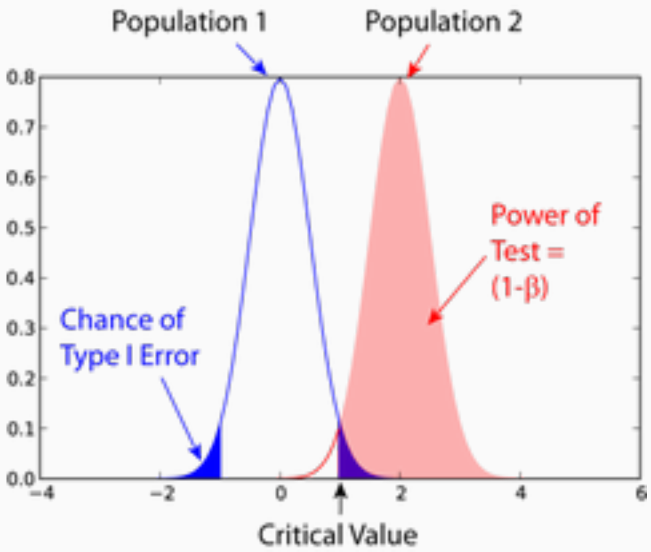
Mainly in relation to probability…

- Probability density curves - integrate to find probability of a value falling in a certain range
- Compute significance of hypothesis tests
- Calculate statistical power of a hypothesis test

```python
import scipy.stats as scs

# instantiate a distribution object with parameters)
dist = scs.# your distribution object here

# this returns the integral of the probability function up to max_val
dist.cdf(max_val)
```



# Partial (multi-variable) Derivatives

Three flavors:

- Gradient (scalar to vector)
- Divergence (vector to scalar)
- Curl (vector to vector)

$$\nabla f = \sum_{x_i} \frac{\partial f}{\partial x_i} \vec{x_i}$$

# Practice

Gradient in cartesian coordinates (x, y, z)

$$\nabla f(x,y,z) = \frac{\partial f}{\partial x}\hat{i} + \frac{\partial f}{\partial y}\hat{j} + \frac{\partial f}{\partial z}\hat{k}$$

Whiteboard example: $f(x,y,z) = 3x^2 + 4xy - 5xyz + 10yz^2$

now you try… $f(x,y,z) = 5x^2 + 4xe^{-y} - x^2y^2z^2 + \frac{1}{3}z^3$

# In real life....

We use linear algebra, so we are taking the gradient of a function involving vectors and matrices

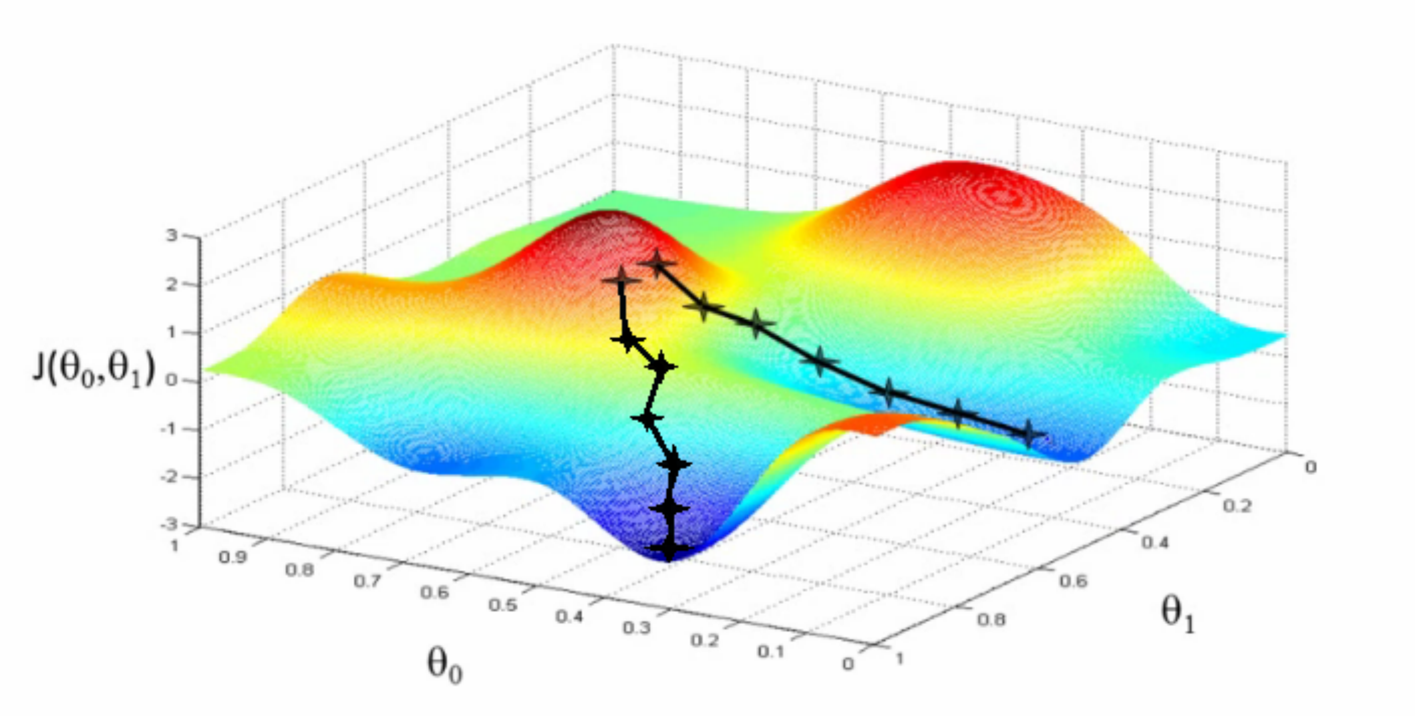
$$\nabla f(\mathbf{X}) = \nabla f(x_1, x_2, ..., x_n) = \frac{\partial f}{\partial x_1}\hat{x}_1 + \frac{\partial f}{\partial x_2}\hat{x}_2 + ... + \frac{\partial f}{\partial x_n}\hat{x}_n$$

Whiteboard example: $f(\mathbf{x}) = f(x_1, x_2) = (a - b\mathbf{x})^2 = (a - b_1 x_1 - b_2 x_2)^2$

# Applications

galvanize

OPTIMIZATION! Try to find the GLOBAL minimum or maximum of a multivariate surface in order to find the best parameters for a function - *gradient descent (and variants)*



# Revisit Learning Objectives

- Review basic calculus operations:
  - single variable derivatives
  - Integrals
  - Partial derivatives / multivariable derivatives
- Motivation : applications of the above

# Resources

https://www.khanacademy.org/math/multivariable-calculus/multivariable-derivatives/partial-derivative-and-gradient-articles/a/the-gradient

https://www.kutasoftware.com/freeica.html