Data Science Math Skills

Katherine G. Pe

Contents

Data Science Math Skills

Course Information

Data Science Math Skills by **Duke University** is a online course you can take on this site Coursera.

Motivation for Learning & Re-learning

A lot of graduate school students struggle with Data Science courses only because of their lack of knowledge and/or understanding of Mathematics for Data Science. The course gives an overview of Mathematical concepts you will encounter while learning Data Science.

Supplemental Notes and Videos

Here's how I make my notes:

My notes include videos from Khan Academy and other websites. The content's the same, and often a bit better due to lack of errors. The text are usually from the Coursera video transcripts.

I indicate **why** it is important to learn such concepts through **Further Reading** notes.

Sets and What They're Good For

Set Basics and Vocabulary

- Set Theory
- Set Theory Operations

Further Reading

> A set is the fundamental discrete structure on which all other discrete structures are built.

Those who studied Discrete Mathematics or read a book about it will probably just re-learn a lot from this course on Set Basics.

- Applications of Set Theory in Computer Science - A list of the most obvious applications of Set Theory.
- Discrete Mathematics and Its Applications - I read most of the book as a supplemental material for a Discrete Math course. The book clearly states why a set is the foundational structure in Computer Science.

Venn Diagrams

- What are Venn diagrams?

Further Reading

- A Visual Explanation of SQL Joins

The Infinite World of Real Numbers

- What are Real Numbers?
- Multi-step Inequalities

The Jagged S Symbol

- Sigma/Summation Notation

$$\sum_{n=1}^{10} n^2$$

- The Sigma has similarities to the factorial symbol, but it suggests that you add the values of i based on the stopping point n

Further Reading

- Graph-based Machine Learning

Descartes Was Really Smart

Plotting Points

- The x-axis is going to be the set of all points x-y in the Cartesian plane, x-y in R2, such that their y coordinate is zero.
- We divide the Cartesian plane into four separate regions, and these we call **quadrants**.
- Coordinate plane: quadrants

Distance Formula

- Distance formula

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Point-Slope Formula for Lines

- Point Slope Form

$$y - b = m(x - a)$$

- Calculating the Slope

$$(y2 - y1/x2 - x1)$$

Slope-Intercept Formula

- Slope Intercept Equation
- Intro to Slope Intercept Form

$$y = mx + b$$

Input-Output Machines

Functions: Mapping from Sets to Sets

- Functions on Sets - UCLA

$$f : A \to B$$

Functions: Graphing in the Cartesian Plane

- Functions on a Coordinate Plane
- Vertical line test
- The **vertical line** test says: *any vertical line, intersects the graph of a function once.* If it intersects it more than once, we violate things here.

Increasing and Decreasing Functions

- A function is increasing whenever:

$$a < b$$

$$f(a) < f(b)$$

$$f(x) = 2^x - \text{An example of an increasing function}$$

- A function is decreasing whenever:

$$a > b$$

$$f(a) > f(b)$$

$$g(x) = 3^{-x} - \text{An example of a decreasing function}$$

Composition and Inverse

- How to find the inverse of a composite function
- Evaluating Composite functions
- Not every function has an inverse.
- If the graph of f fails the *horizontal line test*, the graph has no inverse.

This is about that derivative stuff

Tangent Lines - Slope of a Graph at a Point

- How to find the slope & equation of tangent line at a given point
- Finding the Equation Of A Tangent Line
- A **Tangent Line** is a line which locally touches a curve at one and only one point.
- Formulas:

$$f'(a)$$

Tangent Lines - The Derivative Function

- The derivative & tangent line equations
- Formulas:

$$f'(x) = 2x$$

$$f(x) = x^2$$

Fast Growth, Slow Growth

Using Integer Exponents

- 0 as an exponent:

$$2^0 = 1$$

- Negative exponents:

$$x^{-n} = 1/x^n$$

- Scientific Notation

Simplification Rules for Algebra using Exponents

- Multiplication rule

$$x^n x^m = x^{(n+m)}$$

- Power to a power

$$(x^n)m = x^{nm}$$

- Product to a power

$$(xy)^n = x^n y^n$$

- Fraction to a power

$$\frac{x}{y}^n = \frac{x^n}{y^n}$$

- Division and negative powers

$$\frac{x^n}{x^n} = x^{n-m}$$

- Fractional exponents

$$x^{(\frac{a}{b})} =^b \sqrt{x^a}$$

How Logarithms & Exponents are Related

- Intro to logarithms
- Exponential form:

$$b^x = N$$

- Logarithmic form:

$$x = log_b = N$$

- The logarithm of any base of 1 is 0, just as any number raised to 0 is 1.
- Solving Logarithmic Equations: This video was more helpful in understanding how to solve for x on a Logarithmic equation problem.

General rules

- Product rule

$$log(xy) = log(x) + log(y)$$

- Quotient rule

$$log\frac{x}{y} = log(x) - log(y)$$

- Power and root rule

$$log(x^n) = nlog(x)$$

Further Reading

- One of the most popular concepts in Data Science is in its logarthmic form (ln): Logistic Regression
- Watch more about Simple Logistic Regression

Change of Base

- The base of **10**, **2** and the natural log *(e)* are common in Data Science.
- Formula for changing the base:

$$log_a(b) = \frac{log_x(b)}{log_x(a)}$$

The Rate of Growth of Continuous Processes

- Exponential rate of growth can be a **discrete** exponential rate of growth or a **continuous** exponential rate of growth.
- Note: those who studied Finance will likely see the practicality of knowing, but not of the terms of themselves.
- Discrete rate of growth:

$$1.00(1 + r)t$$

- Euler's contstant e
- What is e?

$$2.71828$$

- Euler is pronounced *Oyler*. He's a mathematician most famously known of his solution to the graph theory problem: Seven Bridges of Königsberg

- Problem: A baby elephant weighing 200 kg grows at a continuously compounded rate of 5%/year. How much does it weigh in 3 years?

$$(200kg)e^{(0.05)(3)} = 232.4kg$$

Basic Probability Definitions

Probability Definitions and Notation

Definition

- Probability is the degree of belief in the truth or falsity of a statement.
- When I am certain that a statement is true then that statement is assigned probability **1** and if I'm certain the statement is false, then it's assigned probability is **0**.

Notation

- The probability of x:

$$P(x)$$

- The negation of statement x:

$$x$$

Law of excluded middle

- Probability of a statement and the probability of the negation of a statement must sum to 1.

$$P(x) + P(\ x) = 0$$

Principle of indifference

- For the *i-th* outcome xi in a distribution with n possible outcomes.

$$P_{xi} = \frac{1}{n}$$

$$P(\text{event}) = \frac{\text{number of outcomes as defined in event}}{\text{total number of possible outcomes in universe}}$$

Further Reading

- Basics of Probability for Data Science

Permutations and Combinations

Permutations

- Order matters
- Example:

$$\text{placing five people in five different positions: 120 ways}$$

- Permutations
- Formula:

$$\frac{n!}{(n-m)!}$$

Combinations

- Order does not matter
- Example:

$$\text{forming a five-person team from five people: 1}$$

- Combinations
- Formula:

$$\frac{n!}{r!(n-r)}$$

Replacement

- Sampling *with replacement* (independent), e.g. drawing a card and putting it back in the deck
- Sampling *without replacement*, e.g. drawing a card from a deck and not putting it back

Using Factorial and "M choose N"

Factorial

- A factorial is the operation where we take a number and multiply it by each integer that is 1 less until we get down to 1.

"m choose n"

- Draw n items from a group of m items without replacement.

$$\binom{m}{n} = \frac{m!}{(m-n)!n!}$$

Further Reading

- This is a recap of Discrete Mathematics. Try to read Discrete Mathematics and Its Applications.

The Sum Rule, Conditional Probability, and the Product Rule

Marginal probabilities and the sum rule

Sum rule

- The marginal probability is equal to the sum of the joint probabilities.

$$P_{(x1)} = P_{(x1,y1)} + P_{(x1,y2)} + P_{(x1,y3)}$$

- Sum rule for binary probability distribution:

$$P(A) = P(A,B) + P(A,\sim B)$$

- Sum rule for series of n probabilities:

$$P(A) = P(A,B1) + P(A,B2) + ... + P(A,B_n)$$

Conditional probability

- The probability that a statement is true given that some other statement is true with certainty.

$$P(A \mid B)$$

Probability of A given that B is true with certainty.

$$P(A \mid B) = \frac{\text{relevant outcomes}}{\text{total outcomes remaining in universe, when B is true}}$$

Product Rule

- Formula:

$$P(A \mid B) = \frac{P(A,B)}{P(B)}$$

Applying Bayes Theorem and Binomial Theorem

Bayes Theorem

- The Bayesian Trap
- Bayes' Theorem and Cancer Screening
- Used for **inverse probability problems**.
- The formula:

$$P(A \mid B) = \frac{P(B \mid A)\,P(A)}{P(B)}$$

- *posterior probability* — probability after new data is observed
- *prior probability* — probability before any data is observed or before new data is observed likelihood— standard forward probability of data given parameters
- *marginal probability* — probability of the data

Inverse Probability

- An inverse probability problem is one where the answer is in the form of the probability that a certain process with a certain probability parameter is being used to generate the observed data.

Binomial Theorem

- Binomial Theorem
- It's binomial because it's used when there are two possible outcomes.
- The formula:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k$$