* every continuous probability distribution has a cumulative density function

↓

which calculate the area up to a given x-value

4) 137 passengers booked on a flight from Vegas to Dallas

40%

Note: the binom is cumulative so need to compute

5) flipped a coin 14 times

15 H

4 T

1 - binom.cdf(5, 15, 4)

~ approx of being 50 or greater

Binomial Distribution

binomial distribution → measures how likely k successes can happen out of n trials given p probability

$\boxed{\text{derive this from scratch at some point}}$

are supy for now

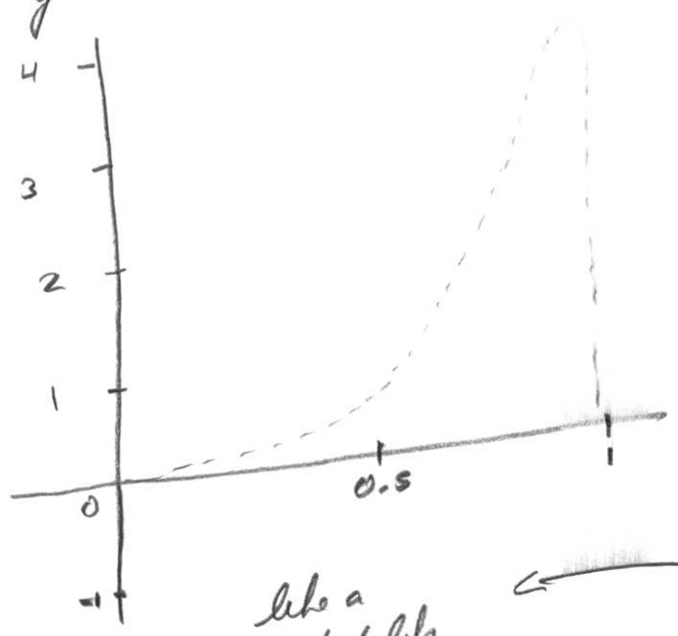
Beta Distribution

what might be problematic about the binomial distribution is I assumed a fixed underlying success rate is 90%.

what about other rates?

instead of creating countless binomial distributions we can use the beta distribution, which allows us to see the likelihood of different underlying probabilities for an event to occur given alpha successes and beta failures

$\boxed{\text{rough sketch}}$



x axis represents all underlying rate of success from 0.0 to 1.0

y axis represent the likelihood of that probability given 8 successes and 2 failures

Beta distribution allows us to see the probabilities of probabilities given 8/10 success

like a meta probability distribution

beta distribution is a continuous function, meaning it forms a continuous curve of decimal values → probability density

↳ analogous to fuct probabilities.

Probability Exercise

1) $P(A) = .3$

$P(B) = .4$

$P(A \text{ and } B) = .3 \times .4 = .012$

2) $P(A \text{ or } B) = .7 + .4 - (.012)$

$.\cancel{9}8$

3) $P(A/B) = \dfrac{P(B/A)\ P(A)}{P(B)}$ and also $P(A \text{ and } B) = P(B) \times P(A/B)$

$P(A \text{ or } B) = P(A) + P(B) - [P(A/B) \times P(B)]$

$P(A) = .3$

$P(B) = .4$

$P(B|A) = .2$

$P(A \text{ and } B) = .4 \times P(A|B)$

$.4 \times \left[\dfrac{P(B|A)\ (P(A))}{P(B)}\right]$

$.4 \times \left[\dfrac{.2\ (.3)}{.4}\right]$

$.4 \left(\dfrac{.06}{.4}\right)$

Essential Math for DS Chapter 2 [Probability] 5

probability is the theoretical study of measuring certainty that an event will happen

$P(x)$ is the probability of event X [decimal]

probabilities refer to the future event state, while likelihoods are measuring the frequency that already occurred.

$$P(\text{not } X) = 1 - (P(x))$$

can also represent probabilities as odds

$$P(X) = \frac{O(X)}{1 + (O(X))} \qquad O(X) = \frac{P(X)}{1 - (P(x))}$$

Probability Math

when we work with a single probability of an event $P(X)$, known as a marginal probability, easy due to mutual exclusivity.

Joint Probabilities

$$P(A \text{ and } B) = P(A) \times P(B)$$

Union Probabilities

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

Conditional Probability and Bayes Theorem

$$P(A \text{ Given } B) \rightarrow P(A|B)$$

Bayes Theorem

$$\boxed{P(A|B) = \frac{P(B/A)\,P(A)}{P(B)}}$$

1) no rational

2) 10^2

3) $81^{1/2} = \sqrt{81} = 9$

4) $25^{3/2} = (25^{1/2})^3$

$(5)^3$

(125)

5) \$1000 = principal

5% interest

12 = monthly

3 yrs = 36 months

$A = P \times \left(1 + \frac{r}{n}\right)^{nt}$

$A = 1000 \times \left(1 + \frac{0.05}{12}\right)^{12 \times 3}$

6) $A = Pe^{rt}$

$A = 1000(e)^{.05 \times 3}$

7) $f(x) = 3x^2 + 1$

$\frac{dy}{dx} = 6x$

8) $f(x) = \int_0^2 3x^2 + 1$

18

$x^3 + x$

$(8 + 2) - (0)$

10

fractional exponent are an alternative way to represent radicals

Lastly what does it mean when we have a fractional exponent with a numerator other than 1.

$(8^{2/3}) = (8^{1/3})^2 = (2)^2 = 4$

cube root then square

(4)

Integrals

For our purposes (again) we will think of the integral. In the Da[illegible]
science, we will be finding the areas under probability distributions.

↳ in this case we will use cumulative density functions that are already integrated. [in appendix A]

$\boxed{\text{Riemann Sum intuition}}$

basically pack rectangles into the area under the curve

$A = P \times e^{rt}$ → is the formula when n becomes continuous.

$\boxed{\text{Limits}}$

sample $f(x) = \frac{1}{x}$, note as x forever increases, $f(x)$ gets closer to 0. Though $f(x)$ never actually reaches 0.

fate of the function, as x goes to infinity, $f(x) \to 0$.

$$\lim_{x \to \infty} \frac{1}{x} = 0$$

we can use SymPy to calculate limits [Note: need to figure out how to represent infinity].

$\boxed{\text{Derivatives}}$

usage of ML derivatives: used in gradient descent and other ML algorithms. [derivative analogy to a slope]

→ $\boxed{\text{partial derivatives}}$ derivative functions that have multiple input variables. For each given variable derivative, we assume the other variables are held constant.

$$f(x,y) = 2x^3 + 3y^3$$

$$\frac{d}{dx} = 3x^2 + 3y^3$$

$$= 3x^2 + 0$$

$$\frac{d}{dy} = 9y^2$$

* we can use SymPy to calculate partial derivatives as well.

$\boxed{\text{The Chain Rule}}$

In chapter 7, when we build a neural network, when we compose the neural network layers, we will have to untangle it (the derivatives) from each layer.

$$y = x^2 + 1$$

$$z = y^3 - 2$$

$$\frac{dz}{dx} = \frac{dz}{dy} \times \frac{dy}{dx}$$

$$y = x^2 + 1$$

$$z = y^3 - 2$$

$$z = (x^2+1)^3 - 2$$

$$z = (x^2+1)^3 - 2$$

$$6x(x^2+1)$$

$$\frac{dz}{dy} = 3y^2 \qquad \frac{dy}{dx} = 2x$$

$$\frac{dz}{dx} = 3y^2(2x)$$

$$= 3(x^2+1)(2x)$$

$$(3x^2+3)(2x)$$

Essential Math for DS Chapter 1

(1)

Number Theory → most of DS will be whole #s, natural #s, integers and real #s.

- natural #s
- whole #s
- Integers
- Rational #s
- Irrational #s
- Real #s
- Complex + Imaginary #s

Imaginary #s → might be used in matrix decomposition

Order of Operations

PEMDAS

Variables

placeholder for unknown variable

Functions

expressions that define relationships between two or more variables

1 takes input [domain variables]
plugs them into an expression
results in an output variable

Summations

$$\sum_{i=1}^{5} 2i = 2(1) + 2(2) + 2(3) + 2(4) + 2(5)$$
$$= 30$$

can be expressed as a for loop

Logarithms

math function that finds a power for a specific number and base.

Euler's # and Natural Logarithms

* property of Euler's number is its exponential function is a derivative to itself which is convenient for exponential and logarithmic functions. In many applications where the base does not really matter, we pick the one that results in the simplest derivative, and that is Euler's #.

Natural logarithms: when we use e as our base for a logarithm, we call it a natural logarithm.

↳ the formula is $A = P \times (1 + \frac{r}{n})^{nt}$

balance = A
starting investment = P
interest rate = r
time span = t
periods = n