Math Camp

September 16, 2017

Unit 2

MSSM Program
Columbia University
Dr. Satyajit Bose

Math Camp

Interlude

Partial Derivatives

- Recall: The derivative $f'(x)$ of a function $f(x)$ is the slope of $f(x)$.
- Alternatively, the derivative $f'(x)$ is the rate at which the function $f(x)$ is changing with respect to changes in the variable x.
- What is the analogue of the derivative in a multivariate function?

Partial Derivatives

- The partial derivative of $f(x,y)$ with respect to x, written $\partial f/ \partial x$, is the derivative of $f(x,y)$ where y is treated as a constant and $f(x,y)$ is considered as if it were a function of x alone.
- The partial derivative of $f(x,y)$ with respect to y, written $\partial f/ \partial y$, is the derivative of $f(x,y)$ where x is treated as a constant and $f(x,y)$ is considered as if it were a function of y alone.

Interpretation of Partial Derivatives

- Since $\partial f/ \partial x$ is simply the ordinary derivative with y held constant, $\partial f/ \partial x$ gives the rate of change of f(x,y) with respect to x for y held constant.
- Alternatively, keeping y constant and increasing x by one (small) unit produces a change in f(x,y) that is approximately given by $\partial f/ \partial x$.

Partial Differentiation

- $f(x,y) = 5x^3y^2$
- Compute $\partial f / \partial x$
  - Think of $f(x,y)$ as $[5y^2]x^3$ where the brackets facilitate treating y as a constant
  - Differentiate with respect to x
- $\partial f / \partial x = 3[5y^2]x^2 = 15x^2y^2$

Partial Differentiation

- $f(x,y) = (4x + 3y - 5)^8$
- Compute $\partial f / \partial x$
  - Think of $f(x,y)$ as $(4x + [3y - 5])^8$ where the brackets facilitate treating y as a constant
  - Differentiate with respect to x
- $\partial f / \partial x = 8(4x + [3y - 5])^7 * 4$

$= 32(4x + [3y - 5])^7$

Example: Heat Loss

- Recall the heat loss function as function of building dimensions.
  - $f(x,y,z) = 11xy + 14yz + 15xz$
- Calculate and interpret $\partial f / \partial x$ if $(x,y,z) = (10,7,5)$
- $\partial f / \partial x = 11y + 15z$
- $\partial f / \partial x$ at $(10,7,5) = 11*7 + 15*5 = 152$
- The quantity $\partial f / \partial x$ is the marginal heat loss with respect to a change in x. With current dimensions (10,7,5), if x were to increase by 1 unit, then the approximate additional heat loss would be 152 units.

Multivariate Optimization

- As with univariate optimization, critical points of a multivariate function are characterized by zero (partial) slopes.

First Order Condition

- For a function of 2 variables $f(x,y)$ that has either a relative maximum or minimum at (x^*,y^*),

$$\partial f(x^*,y^*)/\partial x = 0$$

AND

$$\partial f(x^*,y^*)/\partial y = 0$$

Example: Energy Efficiency

- Recall the heat loss function as function of building dimensions.
  - $f(x,y,z) = 11xy + 14yz + 15xz$
- Suppose we want to design a rectangular building with volume 147,840 cubic feet which has minimal heat loss.
- We must minimize the function $f(x,y,z)$ s.t. $xyz = 147{,}840$

Example: Energy Efficiency

- We must minimize the function $f(x,y,z)$ s.t. $xyz = 147{,}840$
- For simplicity, denote $147{,}840 = V$

$xyz = V$

- Express z in terms of x,y and V

$z = V/xy$

- Substitute into $f(x,y,z)$

$$f(x,y,z(x,y,V)) = 11xy + 14y(V/xy) + 15x(V/xy)$$
$$= 11xy + 14V/x + 15V/y$$

Example: Energy Efficiency

- To minimize this function, we compute partial derivatives w.r.t. x and y; then we equate them to zero to meet the first order condition (ignoring the second order condition for now)

$f(x,y,z(x,y,V)) = 11xy + 14V/x + 15V/y$

- $\partial f/\partial x = 11y - 14V/x^2$
- $\partial f/\partial y = 11x - 15V/y^2$

Example: Energy Efficiency

- Now equate the partial derivatives to zero

$\partial f/\partial x = 11y - 14V/x^2 = 0$

$\partial f/\partial y = 11x - 15V/y^2 = 0$

- These equations give:

$y = 14V/(11x^2)$

$11xy^2 = 15V$

- Substitute for y in the second equation:

$11x(14V/11x^2)^2 = 15V$

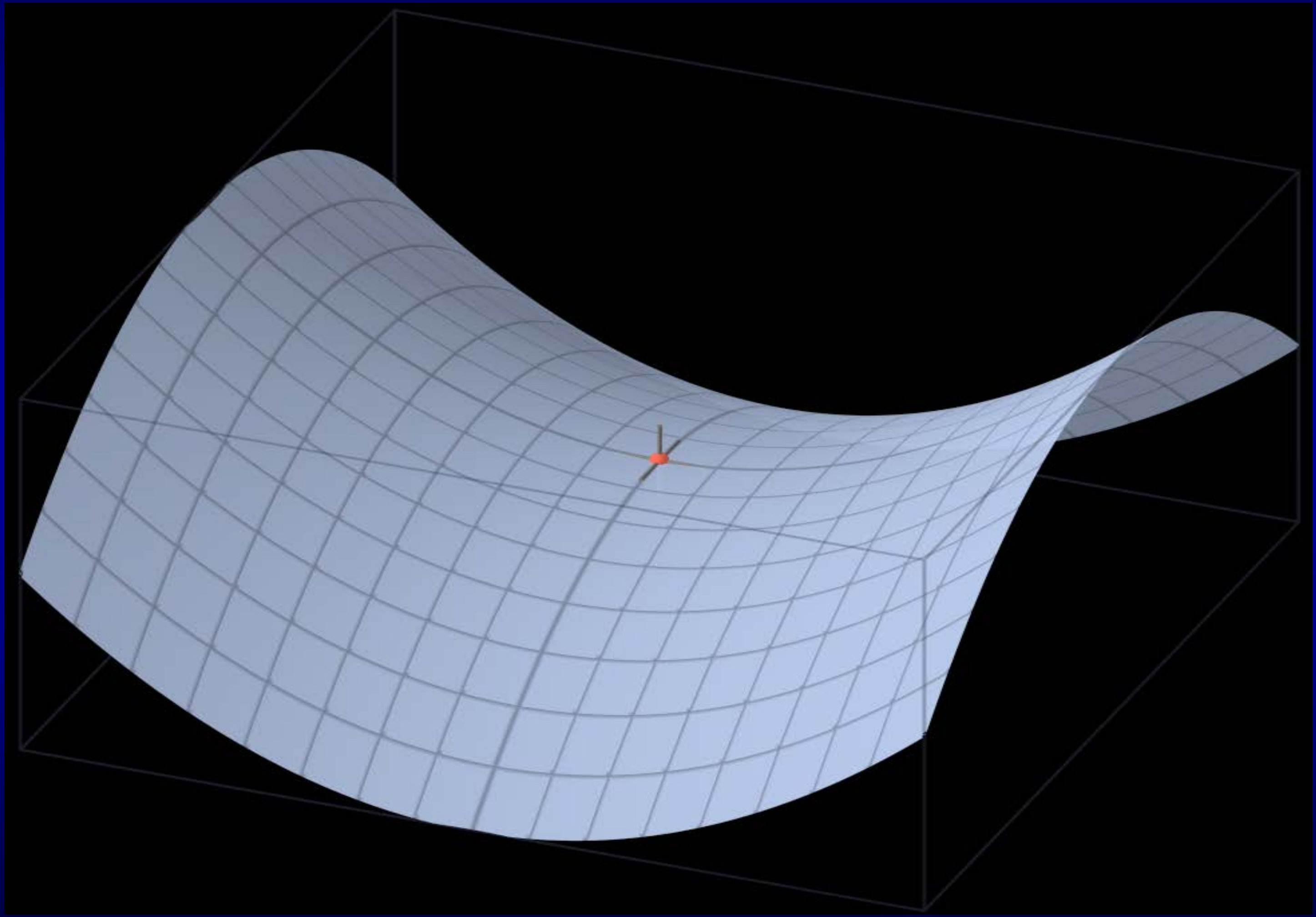
$14^2V^2/11x^3 = 15V$

$x^3 = 14^2V/11*15$

Example: Energy Efficiency

- $x^3 = 14^2V/11*15$
- Using $V = 147{,}840 \Rightarrow x^3 = 175{,}616$ or $x = 56$
- Using $y = 14V/11x^2 \Rightarrow y = 14*147{,}840/11*56^2 = 60$
- Using $z = V/xy \Rightarrow z = 147{,}840/56*60 = 44$
- The building which minimizes heat loss should be 56ft long, 60ft wide and 44ft high.

Saddle Point



Second Order Condition

- For a function of 2 variables $f(x,y)$ s.t. at (x^*,y^*),

$\partial f(x^*,y^*)/\partial x = 0$ and $\partial f(x^*,y^*)/\partial y = 0$

Define $D(x,y) =$

$(\partial^2 f/\partial x^2)(\partial^2 f/\partial y^2) - (\partial^2 f/\partial x\partial y)^2$

Second Order Condition

- If $D(x,y) > 0$ and $(\partial^2 f(x^*,y^*)/\partial x^2) > 0$

Then $f(x,y)$ has a relative minimum at (x^*,y^*)

- If $D(x,y) > 0$ and $(\partial^2 f(x^*,y^*)/\partial x^2) < 0$ Then $f(x,y)$ has a relative maximum at (x^*,y^*)
- If $D(x,y) < 0$

Then $f(x,y)$ has neither a relative maximum nor a relative minimum at (x^*,y^*)

- If $D(x,y) = 0$, then no conclusion can be drawn from this test.

Math Camp

Interlude

Fundamental Calculus Problems

1. Find the slope of a curve at a point.
   $\Rightarrow$ Find the derivative.
2. Find the area of a region under a curve.
   $\Rightarrow$ Approximate the area using a collection of regular shapes whose areas can be computed geometrically.
   $\Rightarrow$ Or find the antiderivative or the integral.

Derivatives

- Recall: The derivative $f'(x)$ of a function $f(x)$ is the slope of $f(x)$.
- Alternatively, the derivative $f'(x)$ is the rate at which the function $f(x)$ is changing with respect to changes in the variable x.
- In many applications, we need to proceed in reverse: we know the rate of change of the function i.e. we know $f'(x)$ but must deduce $f(x)$.

The Antiderivative

- Suppose that $f(x)$ is a given function and $F(x)$ is a function having $f(x)$ as its derivative, i.e. $F'(x) = f(x)$. We call $F(x)$ an integral or an antiderivative of $f(x)$.

Example

Find the antiderivative of $f(x) = x^2$.

Reverse the general power rule to yield:

$F(x)$ could be $1/3x^3$

or $1/3x^3 + 1$

or $1/3x^3 + 5$

or $1/3x^3 + c$, where c is any constant

Example

Find the antiderivative of $f(x) = 2x - (1/x^2)$

Reverse the sum rule and the general power rule to yield:

$F(x)$ is $x^2 + 1/x + c$, where c is any constant

The Indefinite Integral

Suppose that $f(x)$ is a function whose antiderivatives are $F(x) + c$. This is denoted:

$\int f(x)\,dx = F(x) + c$

The symbol $\int$ is the integral sign. The notation $\int f(x)\,dx$ is the indefinite integral. "dx" denotes that the anti-differentiation is with respect to the variable x.

Example

Determine $\int x^r\, dx$ where r is a constant not equal to –1

Note $x^r = ((r+1)/(r+1))\, x^r$

Using the power rule in reverse,

$\int x^r\, dx = (x^{r+1})/(r+1) + c$

$r \neq -1$ since division by 0 is not defined.

What if $r = -1$?

Example

Determine $\int e^{kx}\,dx$ where k is a constant not equal to 0

Note $e^{kx} = (k/k)\, e^{kx}$

Hence,

$$\int e^{kx}\,dx = (1/k)\, e^{kx} + c$$

$k \neq 0$ since division by 0 is not defined.

Integration Rules

1. Sum Rule

   $\int [f(x) + g(x)]\,dx = \int f(x)\,dx + \int g(x)\,dx$

2. Constant Multiple Rule

   $\int kf(x)\,dx = k \int f(x)\,dx$ where k is a constant

Application: Oil Consumption

During the early 1970s, the annual worldwide rate of oil consumption grew exponentially with a growth constant of about .07. At the beginning of 1970, the rate was 16.1 billion barrels annually.

Let R(t) denote the rate of oil consumption at time t, where t = # of years since the beginning of 1970

Application: Oil Consumption

Use the function $R(t) = 16.1e^{0.07t}$ to determine the total amount of oil that would have been consumed from 1970 to 1980 had this rate of consumption continued throughout the decade.

Let $T(t)$ = total oil consumed from 1970 to time t

$T'(t)$ is the rate of oil consumption (by definition of the derivative)

Hence $T'(t) = R(t)$

Hence, $T(t)$ is an antiderivative of $R(t)$

Application: Oil Consumption

$T(t) = \int 16.1e^{0.07t}\, dt$

$= 16.1 \int e^{0.07t}\, dt$

$= 16.1\, ((1/0.07)\, e^{0.07t}) + c$

$= 230e^{0.07t} + c$

We know that $T(0) = 0$

Hence, $230e^{0.07\times 0} + c = 0 \Rightarrow c = -230$

Hence, $T(t) = 230e^{0.07t} - 230$

Application: Oil Consumption

The total amount of oil that would have been consumed from 1970 to 1980 is

$T(10) = 230e^{0.7} - 230 \approx 233$ billion barrels

Application: Marginal Cost

A factory's marginal cost function is

$(1/100)x^2 - 2x + 120$, where x is qty produced per day. The factory has fixed costs of $1000 per day. Find the cost function.

$C'(x) = (1/100)x^2 - 2x + 120$

$C(x) = \int [(1/100)x^2 - 2x + 120]\, dx$

$= (1/300)x^3 - x^2 + 120x + c$

Application: Marginal Cost

$C(0) = 1000$

$\Rightarrow 1000 = (1/300)0^3 - 0^2 + 120(0) + c$

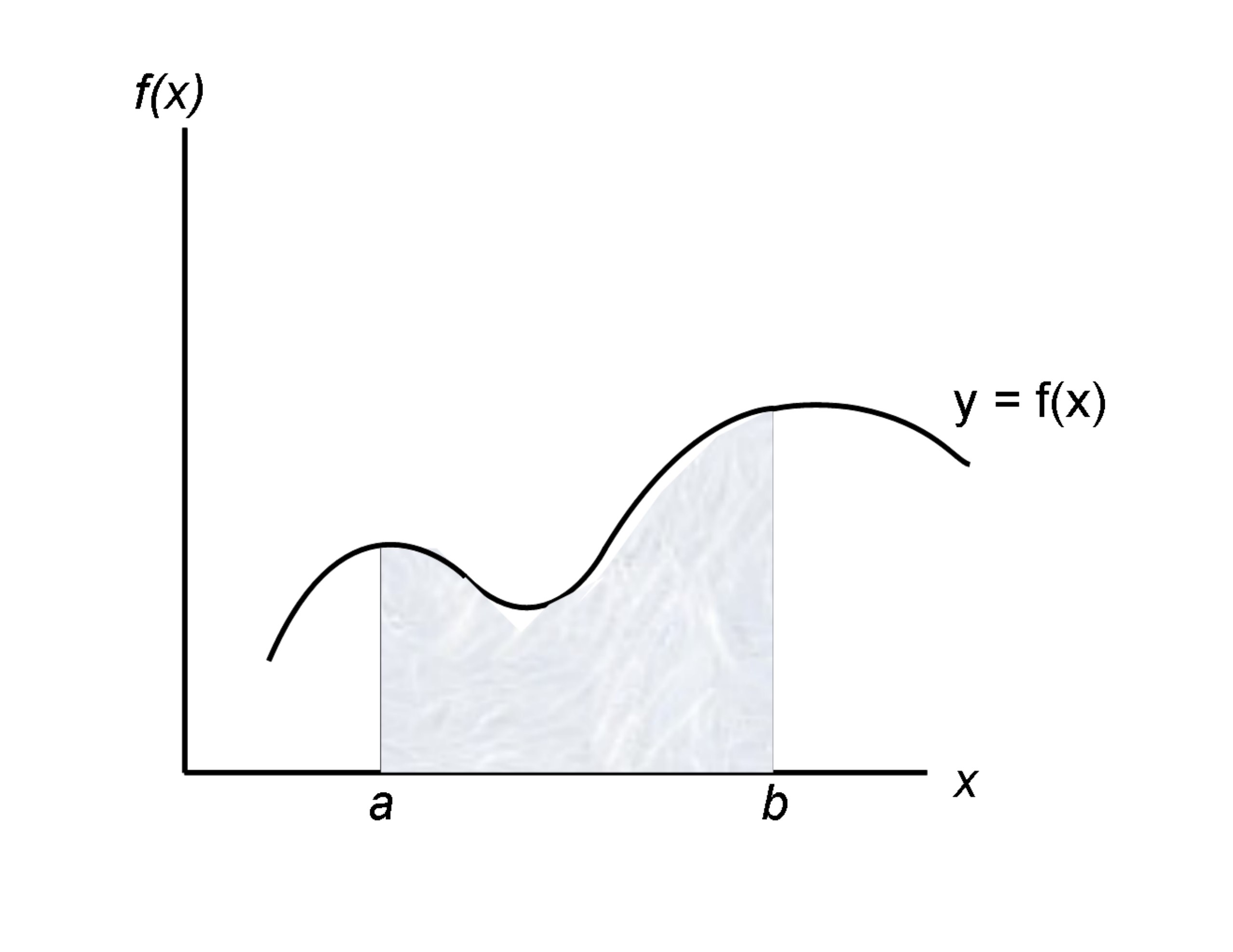
$\Rightarrow c = 1000$

$C(x) = (1/300)x^3 - x^2 + 120x + 1000$

Math Camp

Interlude

Interpretation of the Integral



How do we compute the shaded area under the curve?

Riemann Sum Approximation

Partition the interval [a,b] into n equal sub-intervals, each of width Δx.

For each sub-interval, choose a representative value of x denoted $x_1, x_2, \ldots, x_n$.

The shaded area can be approximated by summing the areas of the n rectangles which approximately cover the area under the curve:

[Area of rectangular approximation] =

$f(x_1)\Delta x + f(x_2)\Delta x + \ldots + f(x_n)\Delta x$

Definite Integrals

The definite integral of f(x) from a to b is denoted by

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

The definite integral is the limit of the Riemann sum as $\Delta x \to 0$.

If f(x) is continuous and nonnegative in [a,b] then the definite integral = area under the graph

Fundamental Theorem of Calculus

Suppose that f(x) is continuous on the interval [a,b] and let F(x) be an antiderivative of f(x). Then

$$\int_a^b f(x)dx = F(b) - F(a)$$

Example

Use the fundamental theorem of calculus to evaluate the following definite integrals:

a) $$\int_1^2 (x)\,dx$$

b) $$\int_0^5 (2x-4)\,dx$$

Application: Oil Consumption

During the early 1970s, the annual worldwide rate of oil consumption was $R(t) = 16.1e^{0.07t}$ where t = # of years since the beginning of 1970

Determine the amount of oil consumed from 1972 to 1974.

Represent this amount as an area.

Application: Oil Consumption

T(4) – T(2) = oil consumed from 1972 to 1974

T′(t) = R(t)

$$T(4)-T(2)=\int_2^4 R(t)\,dt=\int_2^4 16.1e^{0.07t}\,dt$$

$$=\frac{16.1}{0.07}e^{0.07t}\Big|_2^4=230e^{0.07(4)}-230e^{0.07(2)}$$

≈ 39.76 billion barrels

Application: Oil Consumption

T(4) – T(2) can be represented as the area under the curve of y = R(t) in the interval [2,4]

Application: Compounding

An amount of \$25000 is invested in a savings account in which interest is compounded continuously at a rate of 1.9% per year.

a. What is the balance after 1 year?
b. What is the balance after 5 years?

Application: Compounding

For continuously compounding interest, we can use the formula $A=Pe^{rt}$, where A is the future value and P is the present value. In this case, we are solving for A.

What we know:

P=25000 the initial investment

r=0.019 the interest rate

t=1 time (in years)

This gives us   A = 25000e^{(0.019)(1)}   = $25480

Application: Compounding

At an interest rate of 5.8% compounded continuously, when will an investment of \$15000 double itself?

In this example we want to know a time. Again we use the formula $A=Pe^{rt}$, so we are solving for t. What we know:

$r=0.058$

$P=15000$

Since the final amount must be double the principal, we have $A=30000$. Substituting into the formula gives us $30000=15000e^{(0.058)t}$

Application: Compounding

And solving for t yields:

$$\frac{30000}{15000} = \frac{15000e^{0.058t}}{15000}$$

$$2 = e^{0.058t}$$

$$\ln 2 = \ln\left(e^{0.058t}\right)$$

$$\ln 2 = 0.058t$$

$$\frac{\ln 2}{0.058} = t$$

$$t = 11.951$$

Math Camp

Interlude

Probability Definitions

Sample Space: list of all possible outcomes

Event: a set of possible outcomes

Relative Frequency: The relative frequency of event *A* is the proportion of times that event *A* occurs in N trials.

Probability: The probability of event *A* occurring is the relative frequency of event *A* occurring as $N \to \infty$.

Properties of Events

Mutually Exclusive: Events A and B are mutually exclusive if and only if

$P(A \text{ or } B) \equiv P(A \cup B) = P(A) + P(B)$

Collectively Exhaustive: The events A or B or C are collectively exhaustive if

$P(A \text{ or } B \text{ or } C) \equiv P(A \cup B \cup C) = 1$

Partition: A collection of events which are mutually exclusive and collectively exhaustive is known as a partition of the sample space.

Probability Axioms

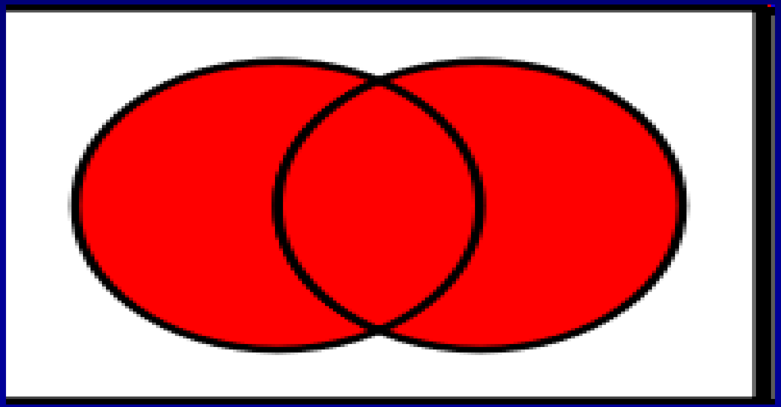
The probability that a discrete valued variable, X, occupies a specific state, x, is a number between zero and one:

$$0 \le P(X = x) \le 1$$

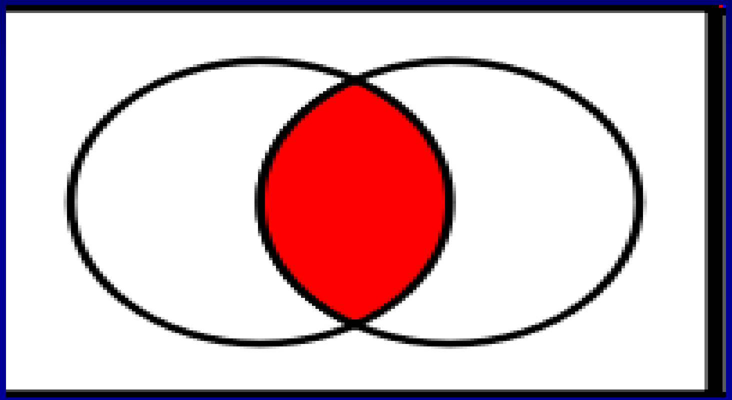
The sum of the probabilities of all outcomes equals one:

$$\sum_{x} P(X = x) = 1$$

Venn Diagrams



Union $A \cup B$

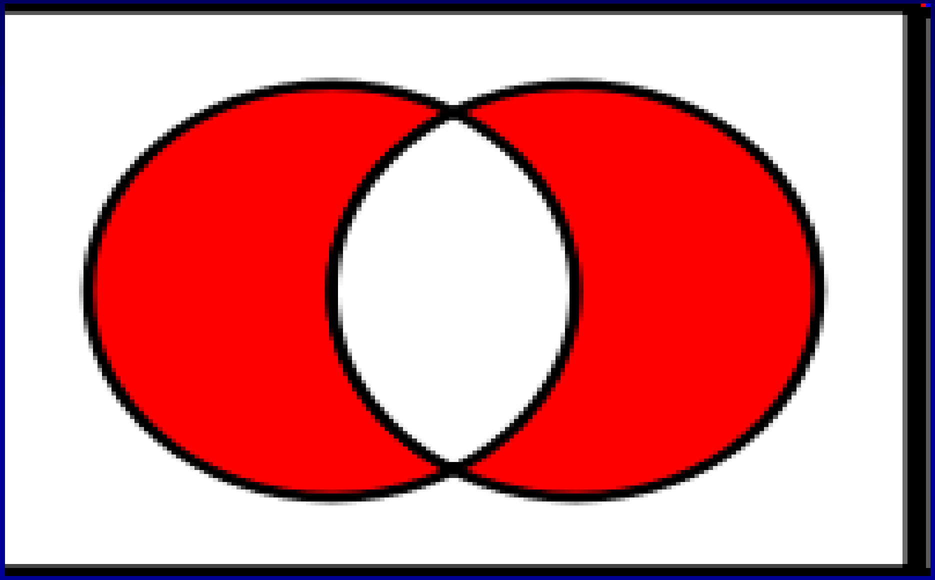


Intersection $A \cap B$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Note: If A, B are mutually exclusive $P(A \cap B) = 0$

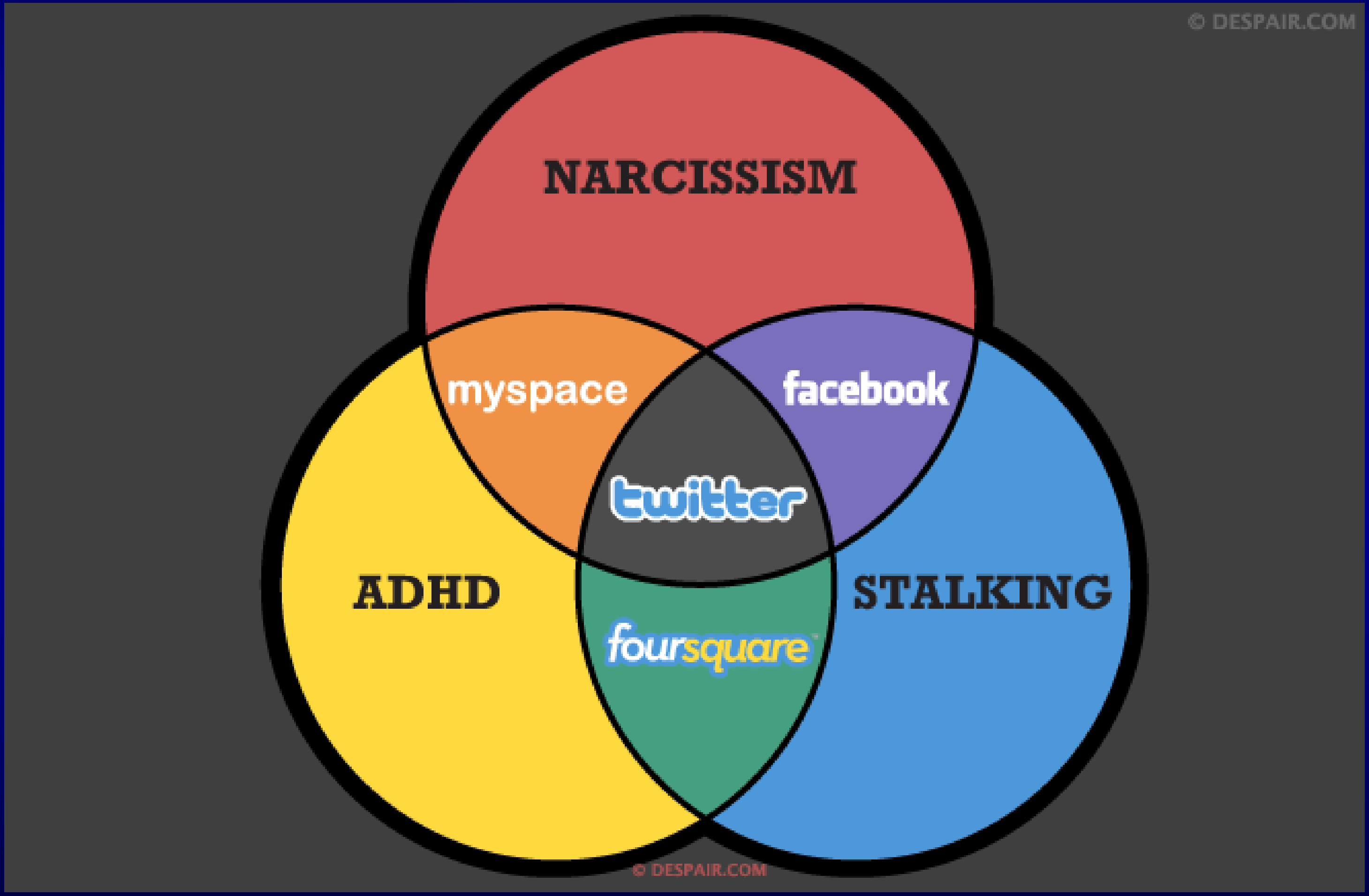
Venn Diagrams



Exclusive Or

$A \oplus B \equiv (A \cup B) \cap (A \cap B)'$

Venn Diagrams



Math Camp

Interlude

Conditional Probability

Often, we want to know the probability of an event given the occurrence of some other event. For example, the probability of an oil tanker spill, given that the captain of the tanker is drunk.

Is the probability of an oil spill given a drunk captain higher or lower than the unconditional probability of an oil spill?

Conditional Probability

The probability of event B occurring conditional on event A equals the probability of events A and B occurring divided by the probability of A occurring:

$P(B|A) = P(A \cap B) / P(A)$

Example

Suppose a die is rolled. Event A is the roll takes a value less than 4. Event B is that the roll is an odd number. What is the probability of the roll being an odd number given that event A has occurred?

Event A = 1, 2, or 3 is rolled.

Event B = 1, 3, or 5 is rolled.

Of the 3 values that comprise event A, 2 are associated with event B. So the probability of event B occurring given event A is 2/3.

Now use the formula. Events A and B occur when a 1 is thrown or when a 3 is thrown. The probability of one of these happening is 1/6 + 1/6 = 1/3.

$P(A \cap B) = 1/3$ $\quad$ $P(A) = \frac{1}{2}$ $\quad$ $\Rightarrow P(B|A) = 1/3 \div \frac{1}{2} = 2/3$

Independent Events

In a casual sense, two events are independent if knowledge that one event has occurred is no cause to adjust the probability of the other event occurring.

Two events, A and B, are independent, if $P(A|B) = P(A)$

Independent Events

If two events, A and B, are independent, then

$P(A \cap B) = P(A)\,P(B)$

Show that this follows from the definitions of conditional probability and independence.

Birthday Problem

In a set of n randomly chosen people, there is a surprisingly high probability that some pair of them will have the same birthday.

E.G.

In a set of 23 randomly chosen people, the probability that some pair of them will have the same birthday is approximately 51%.

Birthday Problem

In a group of 23, there 253 possible pairs to check for the same birthday:

Person 1 could have the same birthday as 22 others, Person 2 could have the same birthday as 21 others, etc.

$22+21+20+19+\ldots+2+1 = 253$

P(A) = probability of a match

P(A') = probability of no match

P(A) = 1 – P(A')

Number each person 1 to 23. Analyze each person in turn. Define events 1 to 23 as the corresponding person not sharing his/her birthday with any of the previously analyzed people .

Birthday Problem

$P(1) = 365/365$

$P(2) = 364/365$

$P(3) = 363/365$ etc.

$P(23) = 343/365$

Due to independence, we can multiply probabilities of these events

$P(A') = P(1) \times P(2) \times P(3) \ldots \times P(23) = (1/365)^{23} \times (365 \times 364 \times 363 \times \ldots \times 343)$

$P(A') = 0.493$

$P(A) = 0.507$

Bayes Theorem

Consider two events A and B. These two events give rise to two other events: not A and not B.

These are denoted A' and B' or $\neg A$ and $\neg B$.

Note that A and A' comprise a partition of the sample space.

Bayes Theorem

How does the probability of B change if we know whether or not A has occurred? In other words, what is the conditional probability $P(B|A)$?

$P(B\,|A) = P(A \cap B) / P(A)$ by definition.

Bayes Theorem expresses the conditional probability $P(B|A)$ using probabilities which may be easier to compute than $P(A \cap B)$ or $P(A)$.

Bayes Theorem

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

Or

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

Example

All people in a population are tested for a medical condition (e.g. measles antibody). Suppose you have the following information about this medical condition and a laboratory test for this condition:

- The chance of a random draw from the population having the medical condition is 0.001
- The chance of a false positive test result from the lab test is 0.01
- The chance of a false negative test result for the lab test is 0.002

Example

Without the test, it is rational to believe that your chance of having the condition is 0.001. If the test comes back positive, what are the chances you have the medical condition?

Event Definition:

A : You have the medical condition.

$\neg A$:You do not have the medical condition.

B : You have a positive test result.

$\neg B$: You have a negative test result.

Example

$P(A) = 0.001$

False positives: $P(B|\neg A) = 0.01$

False negatives: $P(\neg B|A) = 0.002$

Also, compute complements:

$P(\neg A) = 0.999$

$P(\neg B|\neg A) = 0.99$

$P(B|A) = 0.998$

Example

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

$$= \frac{.001\times .998}{.001\times .998 + .999\times .01}$$

$$= .0908$$

Hence, having tested positive, there is 9.08% chance that you have the medical condition.

Math Camp

Interlude

Probability Distributions

Random Variable: is a function that assigns a number to each possible outcome in a chance experiment.

Example: Roll of 2 dice

Sum of dice	2	3	…	7	…	11	12
P	1/36	2/36	…	6/36	…	2/36	1/36

Discrete Probability Distribution

The probability distribution function P(x) is the function that assigns a number to each possible outcome s.t.

$$\sum_{i=1}^{n} P(x_i) = 1$$

Where x_i is the *i*th value that the random variable can take on, $P(\cdot)$ is the probability the random variable takes on a particular value, and n is the number of possible random values.

Continuous Probability Distribution

The continuous probability distribution function f(x) is given by

$$\int_{-\infty}^{\infty} f(x) = 1$$

Where $\int_{a}^{b} f(x)$ is the probability of the random variable, x , falling between the values of a and b , for all a and b .

Cumulative Distribution (Density)Functions

Related to probability distribution functions are cumulative distribution or density functions (CDFs). The density function relate the value of a random variable with the probability of the random variable being less than or equal to that value.

CDF

For a discrete random variable, the CDF is

$$F(x_j) = \sum_{i=1}^{j} P(x_i) \text{ for all } i \leq j$$

For a continuous random variable, it is

$$F(a) = \int_{-\infty}^{\infty} f(x)dx$$

Note the relationship to the definite integral.

Math Camp

Interlude

Statistics

Statistics is concerned with collection, organization and interpretation of numerical data:

- Aims to infer general knowledge from samples
- Explores data to find tentative patterns (hypotheses)
- Tests the validity of hypotheses

Measurement

Objects to be measured in statistics are often coded as :

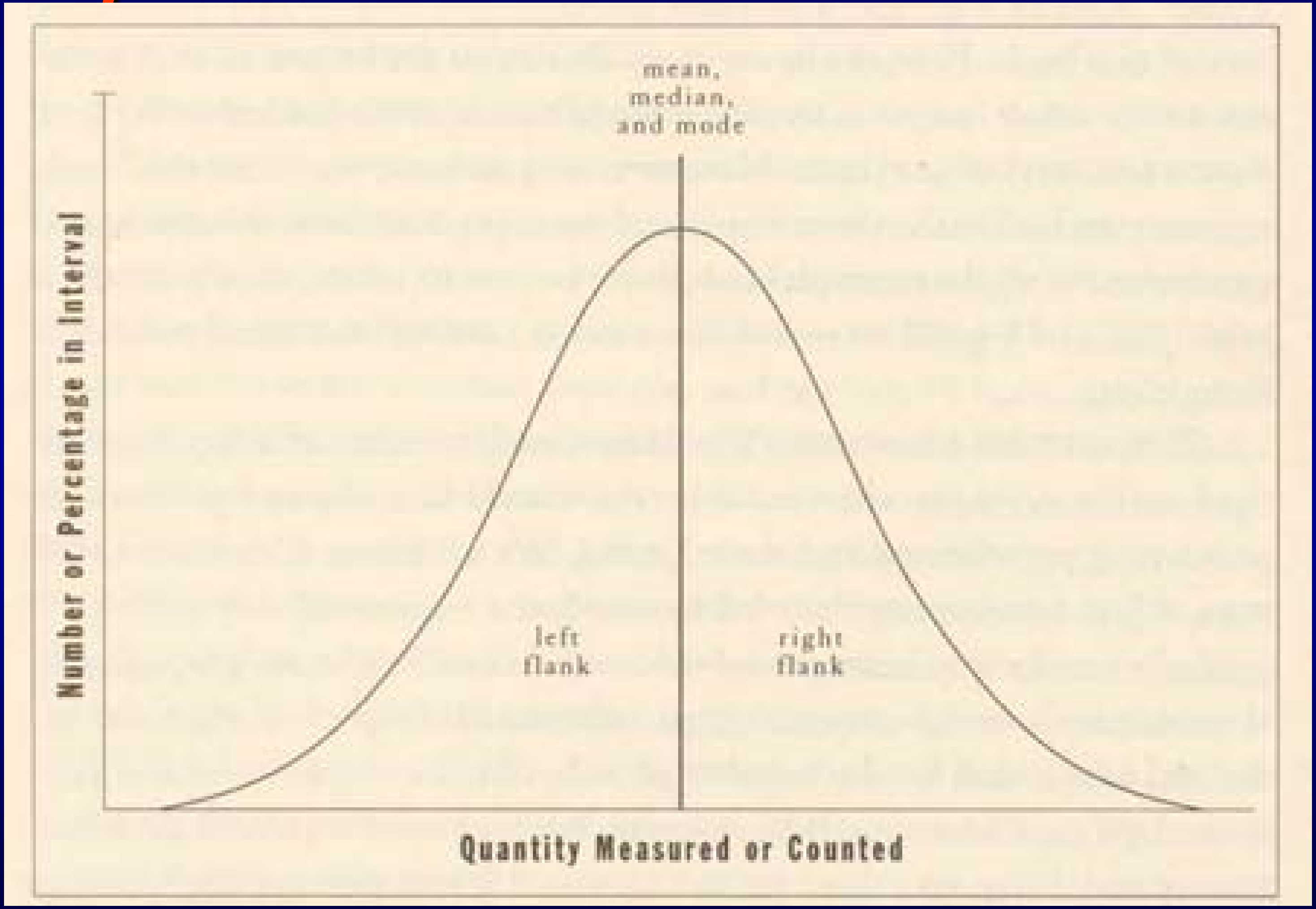
- Nominal variables (qualitiative or categorical)
- Ordinal variables (ranked)
- Scale variables (continuous)

Central Tendency

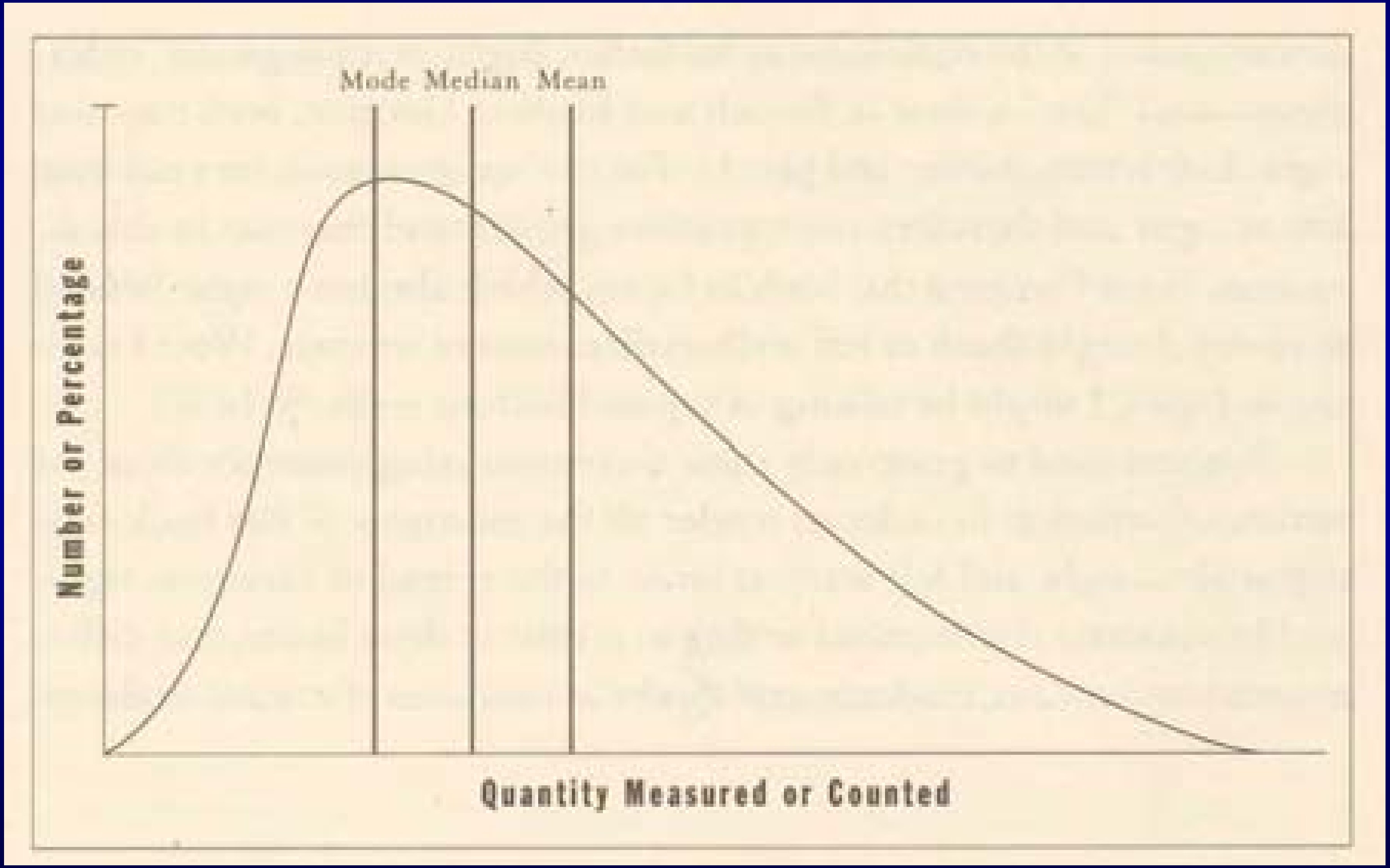
Measures of central tendency:

- Mean = $\frac{1}{N}\sum_{i=1}^{N} x_i$
- Median = lowest value that is ≥ half of the values in the data set.
- Mode = data value with highest frequency

A symmetric distribution



An asymmetric distribution



Dispersion

Measures of dispersion:

- Variance = $\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2 = \sigma^2$
- Standard Deviation = σ