

Hwyl ag ansicrwydd ☺

Philip Jonathan

Prifysgol Caerhirfryn
Ysgol Y Gwyddorau Mathemategol

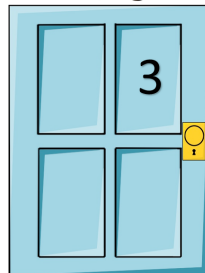
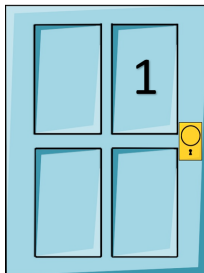
Cynhadledd Y Gymdeithas Ddeintyddol
Hydref 2025

Lancaster
University 

I blant T. Arfon Williams

Heddiw anerchaf ddeintyddion - rwy awr
Am risg yng Ngaernarfon
Wen, â nawdd disgynyddion
Annwyl lais Treherbert lon.

Gafr neu Maserati?

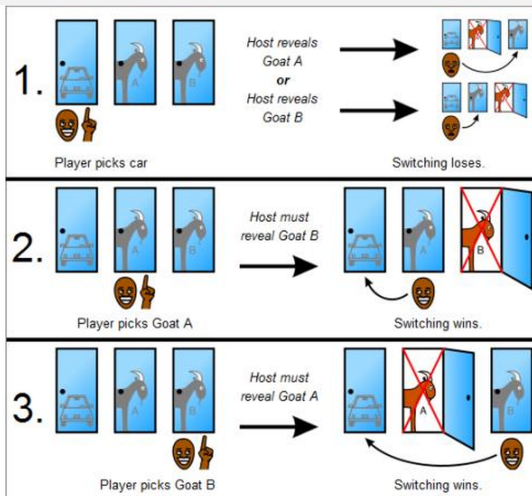


Gafr neu Mazerati?

Dychmygwch ...

- 'Da chi mewn cwis ar y teledu
- Ma' 'da chi ddewis o dri drws
- Tu ôl i un drws mae 'na gar ffansi (rhywbeth **dymunol** iawn!); tu ôl i'r ddau drws arall mae 'na afr (neu rhywbeth ych-a-fi!).
- 'Da chi'n dewis drws, ac yn rhoi rhif y drws i'r boi teledu
- Mae'r boi teledu yn agor un o'r ddau drws **na ddewisoch** sydd â **gafr y tu ôl iddo**
- Mae'r boi teledu'n gofyn **Ydych chi eisiau newid eich dewis?**
- Be' 'da chi'n gwneud? **Ydi e o fantais i chi, i newid eich dewis?**

Gafr neu Mazerati?



... weithiau, doethach newid meddwl

Pryd i briodi?

Dychmygwch

- 'Da chi'n ddyn sengl yn edrych am wraig
- 'Da chi'n mynd i gwrdd â 100 o ferched cymwys ar hap dros nifer fawr o flynyddoedd
- Byddwch chi'n hoffi rhai'n well nag eraill, ac yn hoffi un fwyaf oll!
- Mae'n rhaid i chi benderfynu'n syth pan 'da chi'n cwrdd â merch, os taw hi yw'r un
- 'Da chi ddim eisiau aros am byth i briodi, wrth rheswm!
- Sawl merch dylsech chi gwrdd â nhw cyn penderfynu?
- Cynigion?

Pryd i briodi?

- Mae'r ateb yn esiampl o **reol stopio optimaidd**, un o gonglfeini maes a elwir yn **ymchwil gweithredol**, a dyma fe ateb
- Dylsech chi aros tan i chi gwrdd â $100/e \approx 37$ o ferched a chofio o'r rheini pa un oedd orau (dywedwn ni taw Carys oedd enw hon)
- Wedyn dylsech chi ddal i gwrdd â merched, a dewis fel gwraig y ferch gyntaf 'da chi'n cwrdd â hi sy'n **well na Carys**
- Fel hyn, mae gyda chi'r tebygolrwydd uchaf o gwrdd â'r gorau o'r 100, a'r tebygolrwydd hwnw yw $1/e \approx 0.37$
- e yw cysonyn Euler, $e = 2.7182818...$

Eithafon

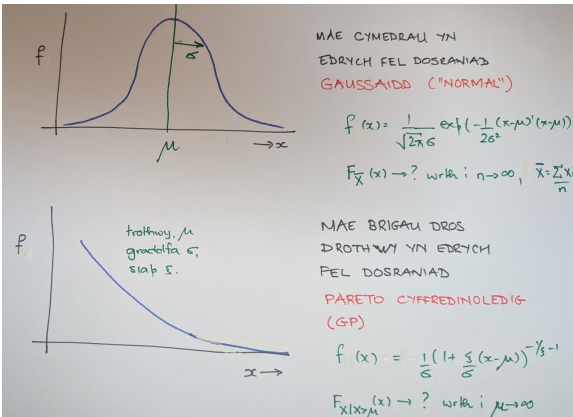
Dychmygwch



- Garwder **storm un-mewn-can mlynedd** ym Môr Y Gogledd yw 10m
- Garwder storm un-mewn-can mlynedd yng Ngwlff Mecsico yw 10m
- Garwder storm un-mewn-**mil** mlynedd ym Môr Y Gogledd yw 13m
- Beth fydd garwder storm un-mewn-mil mlynedd yng Ngwlff Mecsico?

Eithafon

Beth sy'n digwydd pan mae samplau'n mynd yn fawr iawn?



○ Cymedrau → **Gaussaidd**

○ Brigau dros drothwy → **Pareto Cyffredinoledig**

... ac mae'r canlyniadau yn wir **waeth beth** fo dosraniad y data ei hun (o dan amodau cyffredinol iawn)

Eithafon

Garwder storm un-mewn-can mlynedd ym Môr Y Gogledd yw 10m

Garwder storm un-mewn-can mlynedd yng Ngwlff Mecsico yw 10m

Garwder storm un-mewn-mil mlynedd ym Môr Y Gogledd yw 13m

Beth fydd garwder storm un-mewn-mil mlynedd yng Ngwlff Mecsico?

- Mae ffiseg stormydd eithafol yn wahanol ym Môr Y Gogledd a Gwlff Mecsico
- Stormydd **all-drofannol** sydd ym Môr Y Gogledd
- Stormydd **corwynt** sydd yng Ngwlff Mecsico
- Felly, er bod y gwerthoedd un-mewn-can mlynedd yn hafal, nid oes rhaid i'r gwerthoedd un-mewn-mil mlynedd fod
- Mewn gwirionedd, mae'r cynffon Pareto cyffredinoledig yn **hirach** yng Ngwlff Mecsico, and felly bydd garwder storm un-mewn-mil mlynedd yno yn **fwy na 13m**

Cyd-ddigwyddiad?

Dychmygwch

- Eleni tarodd storm enfawr y lan ym Mhenllŷn, storm oedd i'w disgwyl **unwaith mewn can mlynedd**
- Eleni hefyd tarodd storm enfawr y lan yn Aberystwyth, storm oedd hefyd i'w disgwyl **unwaith mewn can mlynedd**
- Beth yw'r tebygolrwydd bod y ddwy storm wedi taro yn yr un flwyddyn?
- Cynigion?

Cyd-ddigwyddiad?

Eleni tarodd storm enfawr y lan ym Mhenllŷn, storm oedd i'w disgwyl **unwaith mewn can mlynedd**. Eleni hefyd tarodd storm enfawr y lan yn Aberystwyth, storm oedd hefyd i'w disgwyl **unwaith mewn can mlynedd**. Beth yw'r tebygolrwydd bod y ddwy storm wedi taro yn yr un flwyddyn?

- Mae'r ateb yn dibynnu ar y **ddibyniaeth** rhwng stormydd yn y ddau le
- Os bod dibyniaeth **berffaith** rhwng y stormydd, y tebygolrwydd yw **$1/100$** : unwaith i'r storm gyntaf daro, anochel yw bod y llall yn taro hefyd
- Os bod y stormydd yn **annibynnol**, y tebygolrwydd yw **$1/100 \times 1/100$** : mae'r ffaith fod y storm gyntaf wedi taro ddim yn dylanwadu ar y siawns bydd yr ail storm yn taro
- Felly mewn gwirionedd gall yr ateb fod unrhywle rhwng **$1/100$ a $1/100 \times 1/100$** , yn dibynnu ar y ddibyniaeth!
- Mae modelu **dibyniaeth** (gofodol, amserol, rhyng-newidynnol) yn rhan annatod o waith ystadegol

Y natur ddynol

Dewisiadau rhifau'r loteri genedlaethol

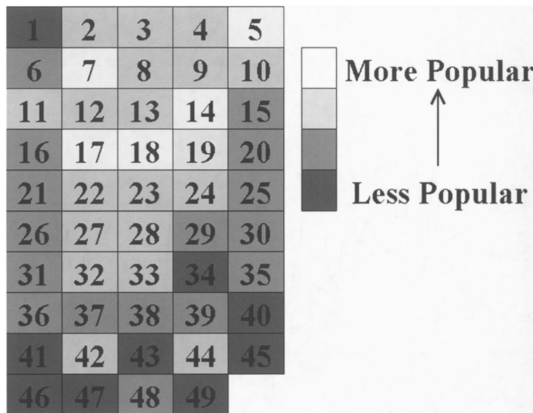
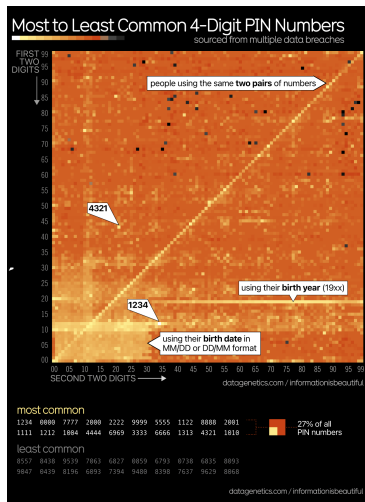


Fig. 2. Estimated probability of players picking each number, laid out as on a lottery ticket: the most popular numbers are unshaded and the least popular are darkly shaded

Ambell waith mae ymddygiad “y dorf” yn rhoi cyfle i eraill!

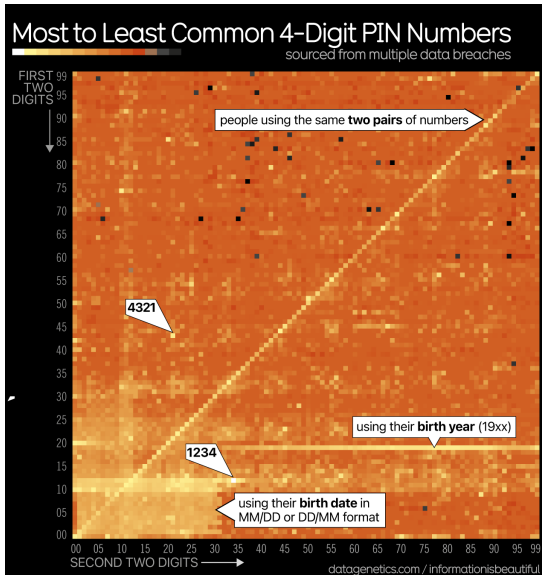
Y natur ddynol

Rhifau pin cyfrifon ar-lein



Ambell waith mae ymddygiad “y dorf” yn rhoi cyfle i eraill!

Y natur ddynol



Dysgu ystadegol

Dychmygwch

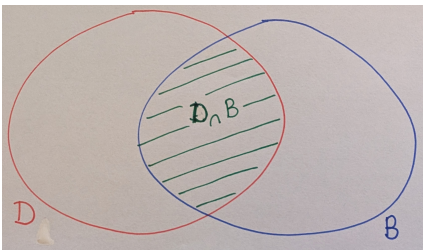
- Mae gennych **gredoau** B a **data** D
- Sut mae gwella'ch credoau yn defnyddio'r wybodaeth yn y data?
- Sut mae dysgu am B gan ddefnyddio D ?
- Oes modd ysgrifennu'r broses i lawr yn fathemategol?

Dysgu ystadegol

Mae gennych greddau B a data D

Sut mae dysgu am B gan ddefnyddio D ?

Meddylwch am setiau



$$\mathbb{P}(D \cap B) \triangleq \mathbb{P}(D|B) \times \mathbb{P}(B)$$

$$\triangleq \mathbb{P}(B|D) \times \mathbb{P}(D)$$

$$\therefore \mathbb{P}(B|D) \times \mathbb{P}(D) = \mathbb{P}(D|B) \times \mathbb{P}(B)$$

$$\mathbb{P}(B|D) = \frac{\mathbb{P}(D|B) \times \mathbb{P}(B)}{\mathbb{P}(D)}$$

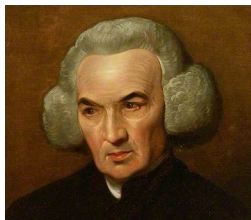
Â dwyseddau

$$f(B|D) = \frac{f(D|B) \times f(B)}{f(D)}$$

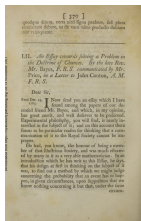
Tebygolrwydd (amodol) a theorem Bayes

$$f(B|D) = \frac{f(D|B) \times f(B)}{f(D)}$$

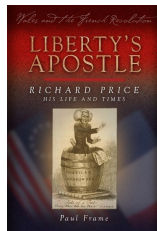
ôl-ddwysedd $\propto$ tebygoliaeth y data $\times$ rhag-ddwysedd
 cred newydd $\propto$ gwybodaeth o'r arbrawf $\times$ cred i gychwyn



Richard Price (Llangeinwyr, 1723 – 1791)



Ei erthygl ar waith Thomas Bayes



Llyfr diweddard amdano

- Rhoi'r debygoliaeth a'r rhag-ddwysedd "fel rhifau" a'u "lluosi"!
- Sail **dysgu ystadegol**

Meintioli risg a dylunio optimaidd

Mae gennym

- System â phriodweddau $\mathcal{S}$ i'w hoptimeiddio
- Mewnbwn X yn adweithio gyda'r system i greu allbwn Y , yn ôl y dwysedd $f(y|x, \mathcal{S})$ sy'n hysbys i ni
- Priodweddau X yn anhysbys, ond gallwn fesur data D
- Mae budd $U(Y|\mathcal{S})$ (neu gost) y system yn hysbys i ni

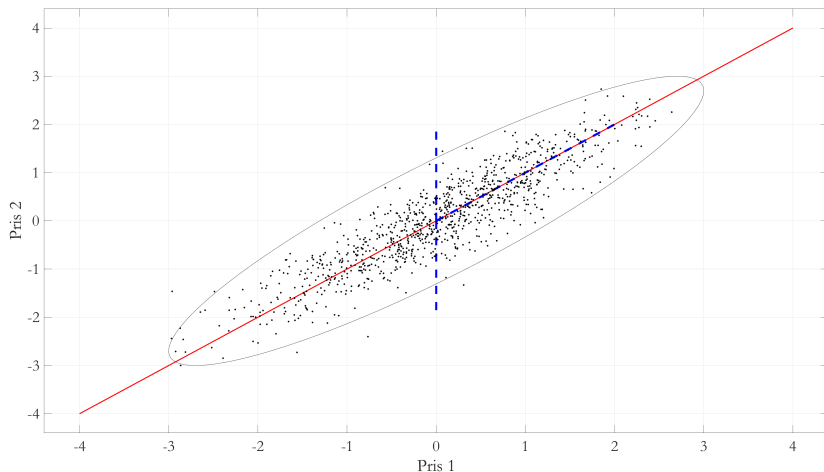
Sut gallwn ddylunio'r system optimaidd?

- Amcangyfrif model Bayesaidd am ddwysedd $f(x|D)$ y mewnbwn
- Gwerthuso'r budd cyfartalog

$$\mathbb{E}[U|\mathcal{S}, D] = \int_y \int_x U(y|\mathcal{S}) f(y|x, \mathcal{S}) f(x|D) dx dy$$

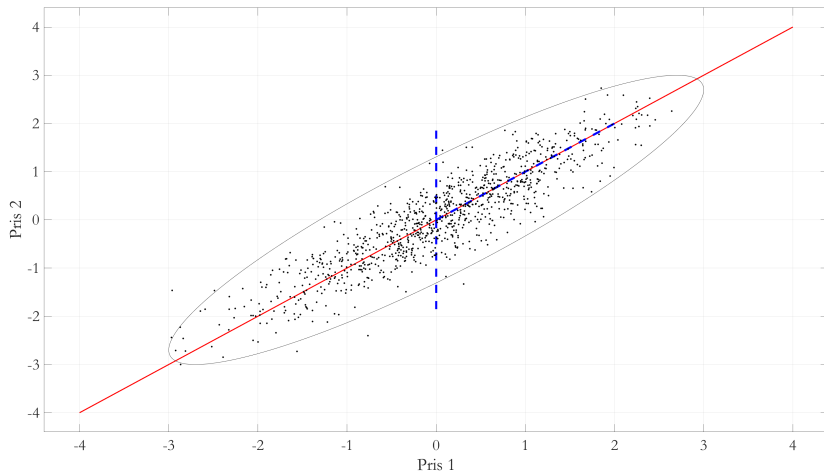
⇒ Gallwn newid $\mathcal{S}$ i uchafsymio'r budd!

Eithafon aml-ddimensiynol Gaussaidd



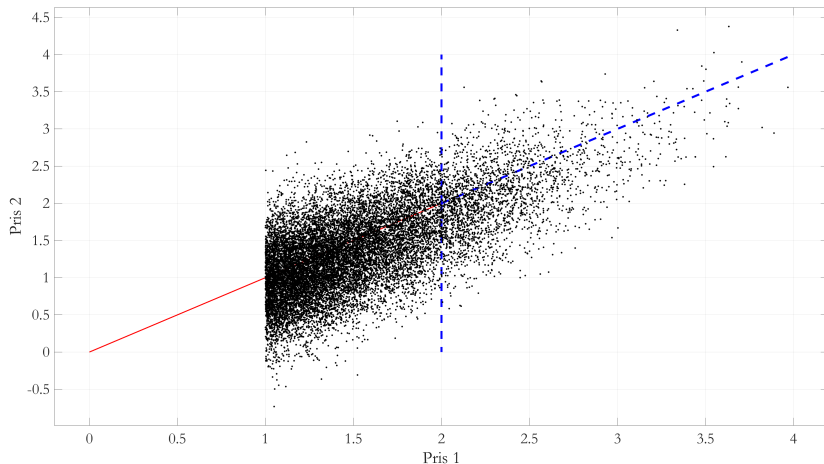
Beth yw'r siawns for Pris 2 yn uwch na Phris 1, pan fydd Pris 1 yn uchel? Beth yw $\mathbb{P}(P_2 > x | P_1 > x)$ pan fydd x yn uchel iawn? Cynigion?

Eithafon aml-ddimensiynol Gaussaidd: cynigion?

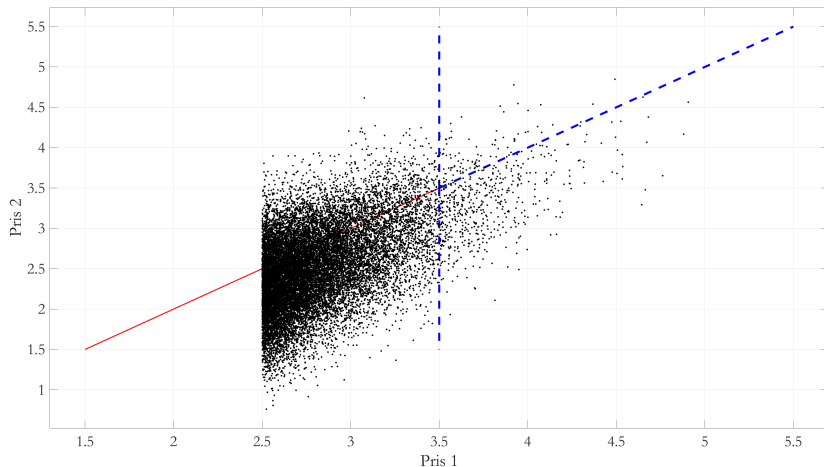


Pan fydd trothwy x Pris 1 yn **isel**, mae $\mathbb{P}(P2 > x | P1 > x) \approx 0.5$

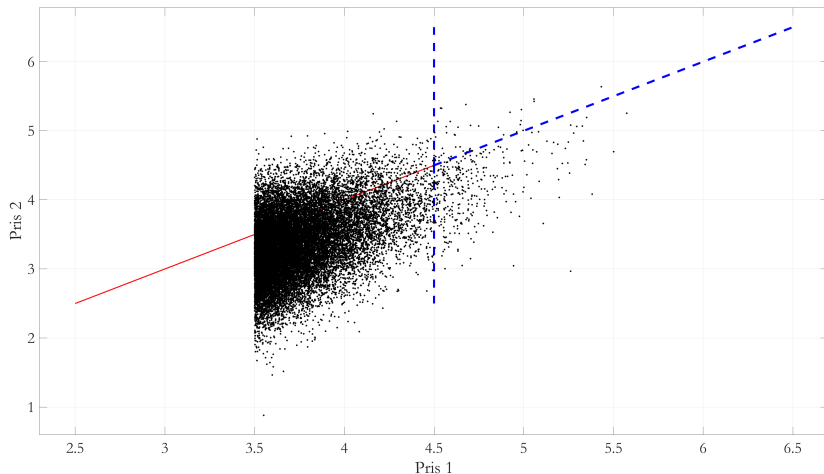
Eithafon aml-ddimensiynol Gaussaidd: Pris 1 > 2



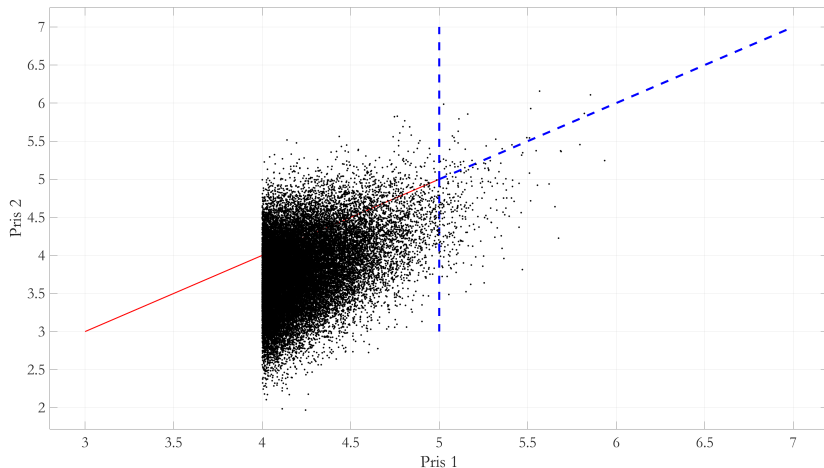
Eithafon aml-ddimensiynol Gaussaidd: Pris 1 > 3.5



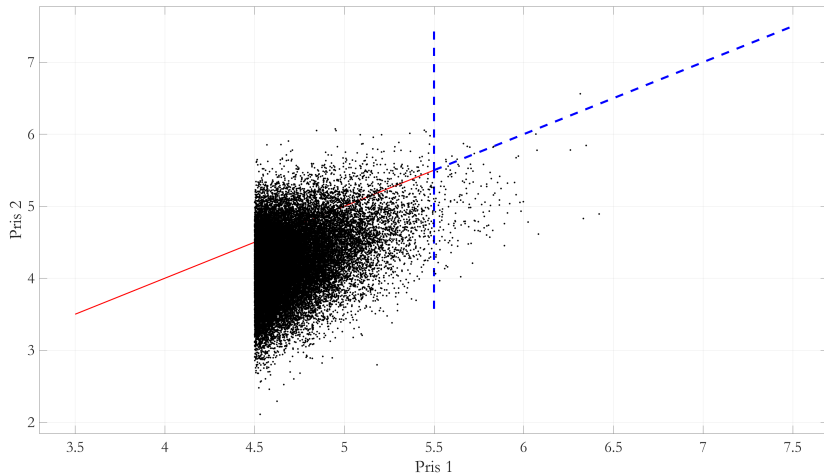
Eithafon aml-ddimensiynol Gaussaidd: Pris 1 > 4.5



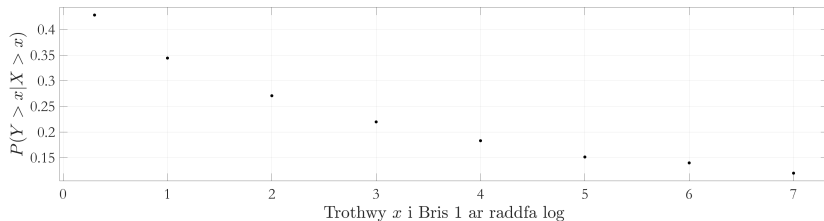
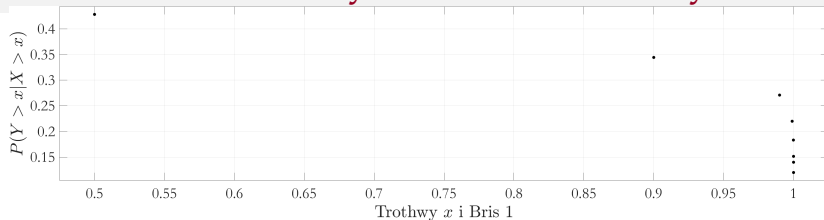
Eithafon aml-ddimensiynol Gaussaidd: Pris 1 > 5



Eithafon aml-ddimensiynol Gaussaidd: Pris 1 > 5.5



Eithafon aml-ddimensiynol Gaussaidd: felly?



Pan fydd trothwy x Pris 1 yn **uchel iawn**, fe welwn fod

$$\mathbb{P}(P2 > x | P1 > x) \rightarrow 0$$

Mae'r canlyniad â goblygiadau mawr ym myd **modelu ariannol**

Maths rhyfedd ffractalau

Dychmygwch

- Rydym yn mynd i iteru'r ymadrodd $z_{n+1} = z_n^2 + c$, $n = 1, 2, \dots$
- Rhifau cymhlyg yw z a c (e.e. $z = z_1 + iz_2$, $c = c_1 + ic_2$, $i = \sqrt{-1}$)
- Amrywiwn werthoedd c_1 a c_2 ac arsylwi os bod $|z_{n+1} - z_n|$ yn mynd yn fawr neu yn fach wrth i n gynyddu
- Crëwn fap, â c_1 a c_2 fel echelinau, a lliwiwn werthoedd sy'n cyd-gyfeirio yn ddu, a'r rhai sy'n dargyfeirio yn wyn
- Pa fath o batrwm wnewn ni greu?

$$z_1 = -0.2 + i0, \text{ ein dewis cychwynnol}$$

$$z_2 = z_1^2 + c$$

$$z_3 = (z_1^2 + c)^2 + c$$

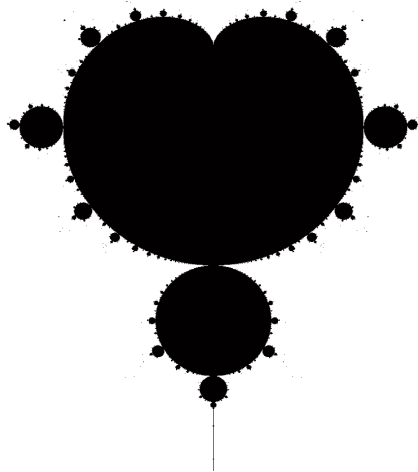
$$z_4 = ((z_1^2 + c)^2 + c)^2 + c$$

$$z_5 = (((z_1^2 + c)^2 + c)^2 + c)^2 + c$$

...

Maths rhyfedd ffractalau

$$z_{n+1} = z_n^2 + c, n = 1, 2, \dots, \text{map ag echelinau } c_1 \text{ a } c_2$$



- Y canlyniad yw set Mandelbrot
- Mae'r set yn **hunan-debyg**
- Mae'r set **bron** yn **gymesur**
- Ar y ffin rhwng du a gwyn, nid yw'r iteru yn cyd-gyfeirio nac yn dargyfeirio! Yma ceir **caos**

Maths rhyfedd ffractalau

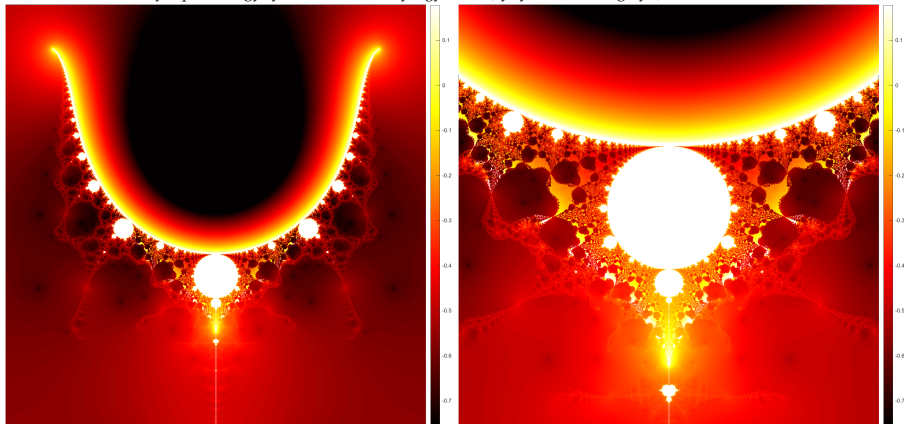


Maths rhyfedd ffractalau



Maths rhyfedd ffractalau

Ffractal **Nova** yn dilyn $z_{n+1} = z_n - (z_n^3 - 1)/(3z_n^2) + c$, darganfyddwyd (dyfeisiwyd?) gan Paul Derbyshire yn yr 1990au. Mae'r lliwiau'n dweud wrthym pa more gyflym mae'r iteru'n cyd-gyfeirio. (cyflym=du, araf=gwyn)



Mae iteriad Nova yn cynnwys iteriad **dull Newton** $x_{n+1} = x_n - \frac{f(x)}{f'(x)}$ i ddarganfod gwerthoedd x sy'n rhoi $f(x) = 0$ am unrhyw ffwythiant f

Maths rhyfedd **caos**

Dychmygwch

- Rydym yn mynd i ddatrys y system o hafaliad differol

$$\frac{dx}{dt} = a(y - x)$$

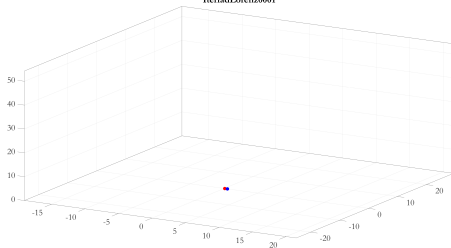
$$\frac{dy}{dt} = x(b - z) - y$$

$$\frac{dz}{dt} = xy - cz$$

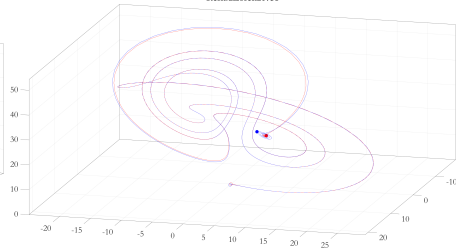
- Mae'r hafaliadau yn dweud wrthym sut y bydd gronyn yn symud mewn amser (t) ar hyd echelinau x , y a z , gyda'r gwerthoedd $a = 10$, $b = 28$ a $c = 8/3$
- Crëwn fap, ag echelinau x , y , z i weld sut bydd **dau** ronyn sy'n dechrau'n agos at ei gilydd yn symud
- Beth sy'n mynd i ddigwydd?** (... a mynd at y fideo)

Maths rhyfedd caos

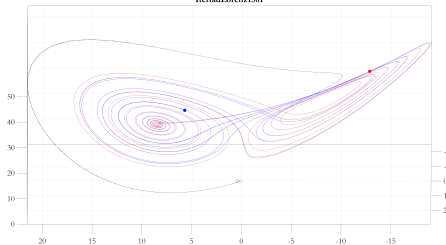
IteriadLorenz0001



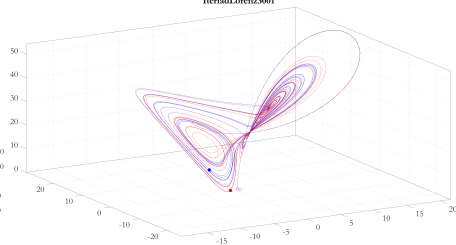
IteriadLorenz0751



IteriadLorenz1501



IteriadLorenz3001



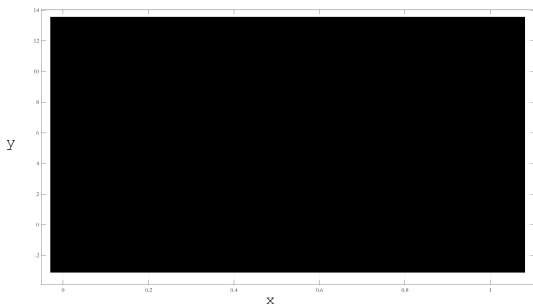
Maths rhyfedd **caos**

- Os gwyddwn lleoliad y gronyn **yn union** (heb unrhyw wall o gwbl) ar dechrau'r daith, fe allwn rhagdybio'i leoliad yn y dyfodol yn berffaith
- Os bod hyn yn oed yr ansicrwydd lleiaf yn ein gwybodaeth am leoliad y gronyn ar y dechrau, fe allwn rhagdybio am **gyfnod yn unig**, wedyn mae taith y gronyn yn mynd yn anhysbys i ni
- Nid yw'r ddau ronyn yn mynd yn **rhy** bell oddiwrth ei gilydd
- Mae ymddangosiad caos yn ddibynnol ar ein dewis o gysonion a , b a c .
- Siâp enwog pili-pala sydd i'r atynnydd (10, 28, 8/3)
- Heddiw defnyddir **ensemble o ragdybiaethau** ar gyfer rhagolwg tywydd, ag aelodau'r ensemble yn dechrau o dan amodau atmosfferig "sy'n agos" at ei gilydd. Os bod rhagdybiaethau'r ensemble yn **aros yn debyg dros amser**, gallwn fod yn fwy hyderus ohonynt.

Gwyddor data: rhyngosod, allosod a dylunio arbrawf

Dychmygwch

- Tu nôl i'r blwch du, mae 'na gromlin $y = f(x)$. Rydym eisiau dysgu amdani
- Gallwn ddewis 10 pwynt ar echelyn x i wneud mesuriad o y
- Ble dylem ni fesur?

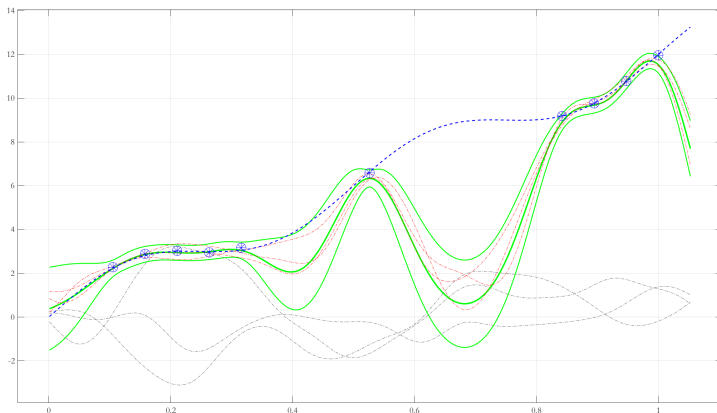


Gwyddor data: rhyngosod, allosod a dylunio arbrawf

Glas = gwir; Gwyrdd = amcangyfrif o gasgliad Bayesaidd

$y = \mathcal{GP}(x|\lambda)$, model **proses Gaussaidd** â hyper-baramedr λ **sydd yn rhy fach**

Llwyd = efelychiadau o dan y rhag-fodel; Coch = efelychiadau o dan yr ôl-fodel



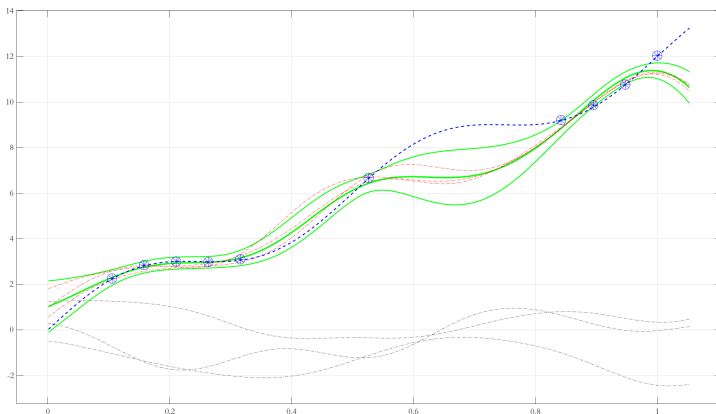
Rhy arw ... mae'r model yn rhyngosod ac allosod yn wael

Gwyddor data: rhyngosod, allosod a dylunio arbrawf

Glas = gwir; Gwyrdd = amcangyfrif o gasgliad Bayesaidd

$y = \mathcal{GP}(x|\lambda)$, model **proses Gaussaidd** â hyper-baramedr λ **sydd yn rhesymol**

Llwyd = efelychiadau o dan y rhag-fodel; Coch = efelychiadau o dan yr ôl-fodel



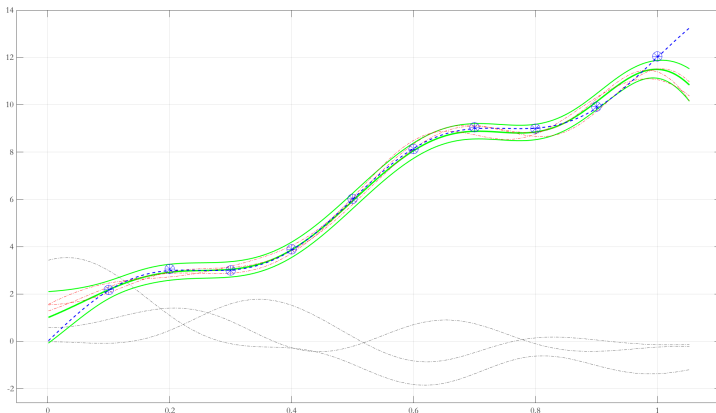
Mwy stiff ... ac yn rhyngosod yn well

Gwyddor data: rhyngosod, allosod a dylunio arbrawf

Glas = gwir; Gwyrdd = amcangyfrif o gasgliad Bayesaidd

$y = \mathcal{GP}(x|\lambda)$, model **proses Gaussaidd** â dewisiadau ar gyfer x sydd yn **gofod-lenwi**

Llwyd = efelychiadau o dan y rhag-fodel; Coch = efelychiadau o dan yr ôl-fodel



Y gorau ... rydym wedi **dylunio'r arbrawf** drwy **ofod-lenwi** i optimeiddio'r model

Gwyddor data: dysgu peirianyddol (ML) a deallusrwydd artiffisial (AI)

Paham fod gymaint o ffÿs a ffwdan am ML ac AI?

- Google AI Overviews, Google Translate
- ChatGPT, Grok, OpenAI, Llama, Claude, Meta AI

Gwyddor data: dysgu peirianyddol (ML) a deallusrwydd artifisial (AI)

Paham fod gymaint o ffÿs a ffwdan am ML ac AI?

- **Cyd-gyfeiriad meysydd gwahanol**
- Casglu data ar raddfa **anhygoel**
- Bron popeth (mae'n ymddangos) “ar-lein” **wedi ei ddigido**
- **Rhwydweithiau** cyd-gysylltiedig, cyflym, byd-eang
- **Gwellianau cyfrifiadurol** (cyflymder, cof, cost, “y cwmwl”)
- **Algorithmau slic** i amcangyfrif modelau empirig
- Y **rhyngwyneb** rhwng dyn a chyfrifiadur yn aeddfedu

Yn ymarferol

- Mynediad hawdd i bob data o ddiddordeb o bobman
- Mynediad hawdd i feddalwedd ystadegol safonol hawdd-ei-defnyddio rhad-ac-am-ddim (e.e. PYTHON, R)
- Mynediad hawdd i gyfrifiaduron pwerus

Gwyddor data: dysgu peirianyddol (ML) a deallusrwydd artifisial (AI)

Ond ...

- Mewn 10^{10} dimensiwn, a ydym yn **allosod** drwy'r amser?
- **Dylunio arbrawf** yn hanfodol, ond **sgrafellu** sy'n digwydd
- Asesu **ansawdd rhagfynegiadau** dylai fod yn hanfodol
- Mae'n faes hynod gyffrous, ond **anaeddfed** (e.e. *"double descent"*, *"stochastic gradient descent"*)
- Nid yw rheolau mathemateg wedi newid!

Gwyddor data: faint o ddata sydd ei angen ar LLM?

$$10^{10} = 10,000,000,000$$

Gwyddor data: faint o ddata sydd ei angen ar LLM?

$$10^{10} = 10,000,000,000$$

Y nifer o ddimensiynnau mewn model LLM
(mewn egwyddor, o leiaf)

Gwyddor data: faint o ddata sydd ei angen ar LLM?

$$2^{10^{10}} \approx 10^{4 \times 10^9}$$

("Un â 4,000,000,000 sero ar ei ol!")

Gwyddor data: faint o ddata sydd ei angen ar LLM?

$$2^{10^{10}} \approx 10^{4 \times 10^9}$$

Y nifer o arsylwadau sydd eu hangen i wahaniaethu rhwng
“isel” ag “uchel” mewn 10^{10} dimensiwm

Gwyddor data: faint o ddata sydd ei angen ar LLM?

10^{16} 10^{16} 10^{24} 10^{90}

Gwyddor data: faint o ddata sydd ei angen ar LLM?

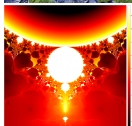
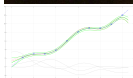
- 10^{16} : nifer o ffeiliau ar y rhwynggrwyd
- 10^{16} : nifer o luniau (h.y. “delwedddau”) yn y byd
- 10^{24} : nifer o feitiau o ddata ar y rhwynggrwyd
- 10^{90} : nifer o ronynnau sylfaenol (e.e. cwarciau, electronau, ffotonau) yn y **bydysawd gweladwy**

Gwyddor data: faint o ddata sydd ei angen ar LLM?

- 10^{16} : nifer o ffeiliau ar y rhyngrwyd
- 10^{16} : nifer o luniau (h.y. “delwedddau”) yn y byd
- 10^{24} : nifer o feitiau o ddata ar y rhyngrwyd
- 10^{90} : nifer o ronynnau sylfaenol (e.e. cwarciau, electronau, ffotonau) yn y **bydysawd gweladwy**
- DS: cofier am gyfuniadau a thrynewidion
- $10^4 \times 10^9$: nifer o arsylwadau i hyfforddi'r LLM yn dda
- Gofod LLM yn **dyllau i gyd** (fel ein bydysawd yn wagle i gyd)
- Dimensiwn **ymarferol** LLM llawer yn is na 10^{10}

Crynodeb

Fe glywsoch am rhain



- Tebygoliaeth amodol, geifr a cheir coch cyflym
- Pryd i briodi, a rheolau stopio optimaidd
- Dibyniaeth ym mhobman
- Dysgu ystadegol, Bayes a Richard Price
- Eithafon a chynffon Pareto cyffredinoledig
- Diffygion y natur ddynol
- Pethau sy'n edrych fel petaent ar hap
- Peryglon allosod a modelau gor-arw
- Da a drwg gwyddor data, ML ac AI

... gwelwch chi, mae ansicrwydd yn gallu bod yn hwyl ☺

Wrth gefn

Gwyddor data: rhyngosod, allosod a dylunio arbrawf

YN FRAS, FAINT O DDATA SYDD ANGEN I AMCANGYFRIF LLM FEL CHATGPT?

I amcangyfrif model syml, mae angen 2 arsylwad mewn pob dimensiwn: 2 arswylad yn 1D, a 4 arsylwad yn 2D ... a 2^p mewn pD

Yn ôl y sôn, mae gan LLM hyd at $10^{10}=10,000,000,000$ paramedr neu ddimensiwn

Felly mae angen hyd at $2^{10,000,000,000}$ arsylwad i amcangyfrif y model yn gall mewn egwyddor

YN FRAS, FAINT O ARSYLWADAU SYDD AR GAEL MEWN GWIRIONEDD?

Y RHYNGRWDYD: Yn 2025, amcangyfrifir mai maint y rhyngrwyd gyfan yw 10^{24} beit. Gan gymryd bod un beit yn arsylwad ...

Gan gofio bod $2^8 = 10^3$ mwy neu lai, mae gan y rhyngrwyd $10^{24} = 2^{64}$ arsylwad. Mae hwn tipyn yn llai na sydd angen!

Y BYDYSAWD GWELADWY: Mae rhyw 10^{82} o ronynnau yn y bydysawd gweladwy

(Oed y bydysawd y 10^{10} mlynedd. 1 blwyddyn golau yw 10^{15} m. Felly radiws y bydysawd gweladwy yw 10^{26} m. Gan gyfrif ehangiad y bydysawd hefyd, coda i 5×10^{26} m. Gan gymryd bod y bydysawd yn spheraidd, ei gyfaint felly yw 5×10^{80} m. Gan gymryd bod yna fwyafswm o 20 gronyn sylfaenol (fermion, boson) pob m^3 , y nifer o ronynnau yn y bydysawd gweladwy yw 10^{82}). Gan gymryd bod un gronyn yn arsylwad ...

Felly mae'r bydysawd cyfan yn cynnig $10^{82} = 2^{8 \times 82/3} = 2^{217}$ o arsylwadau. Mae hwn tipyn yn llai na sydd angen!

YN FRAS, BETH YW ARWYDDOCAD HYN I GYD?

Mae gofod yr LLM yn dyllau heb wybodaeth i gyd ag ambell ynys o arsylwadau (yn debyg iawn i'n bydysawd sy'n ofod i gyd ag ambell seren a phlaned)

Felly mae'r model LLM yn **gorfod rhyngosod** trwy'r tyllau heb wybodaeth. Ni all hyn fod bob amser yn gall, yn amlwg

Mae dimensiwn ymarferol LLM tipyn yn is na 10,000,000,000

Ond: cofier cyfuniadau a thrynewidion o foltiau.