Lecture 9
Understanding Jets

SDSS

SDSS
13.8 Lookback time [Gyrs]
13.5
12.0
10.0
5.0
0.0
eBOSS + BOSS Lyman-α (2008-2019)
eBOSS + SDSS I-II Quasars (1998-2019)
eBOSS Young Blue Galaxies (2014-2019)
eBOSS Old Red Galaxies (2014-2019)
BOSS Old Red Galaxies (2008-2014)
SDSS I-II Nearby Galaxies (1998-2008)
BAO feature strength
50
100
150

Laboratory

- The challenge of this lab is to fin the W and Z boson

Laboratory

- The challenge of this lab is to fin the W and Z boson
  - These are the particles that cover the Weak Boson Interaction

Fundamental Forces

Weak Interactions

- The Weak interaction is described by W and Z bosons

Weak Interactions

- The Weak interaction is described by W and Z bosons

What really is this?

- Weak intreaction is actually a very strong interaction
  - Its just mediated by a heavy particle

Heavy W boson particle produced as an intermeidary

All the Fundamental Particles

W boson Properties

- W boson interacts with many different particles

W boson Properties

- W boson interacts with many different particles
  - These don't happen frequently b/c W boson is heavy

Electroweak Bosons

- Weak force is combined with EM forces to be electroweak
  - Practically, from low to high energy new forces turn on

Z boson Properties

- Z boson is like a heavy photon (no charge, but heavy)
- Can decay to leptons

Z boson Properties

Decays to all quarks, except top

- Z boson is like a heavy photon (no charge, but heavy)

The Forces

- The challenge of this lab is to fin the W and Z boson
  - These are the particles that cover the Weak Boson Interaction

All Part of Electroweak Force

what about Strong force?

Strong Force

- Strong force is what bins quarks together
  - This is how we get protons

Strong Force Production

- Strong force
  - At high energy quarks can get produced which radiate gluons

Understanding Jets

Understanding Jets(modern view)

How does this cluster appear?

quark

How does this cluster appear?

How does this cluster appear?

How does this cluster appear?

A fractal

- The repeated branching structure gives a fractal

Branches

Norway coastline

Romensco cauliflower

Lets use this property to make a simple observable?

What is the scaling dimension of a jet

Box Counting(Hausdorff Dimension)

For a given distance scale :

What number of boxes are needed to describe a jet

Same jet from 2 slides ago

$$\text{(Fractal) Dimension} = \lim_{D(\text{istance})\to 0} \frac{\text{Log}(N_d^{\text{boxes}})}{\text{Log}(\text{Distance})}$$

Project with CERN summer student 2014

Can we characterize this?

- Measuring fractals:

Coastline of Great Britain is D=1.25

$$\text{Dimension} = \lim_{D \to 0} \frac{\text{Log}(N_D^{\text{boxes}})}{\text{Log}(D)}$$

Box Counting

1 Box with distance=1 covers all points

Number of boxes to cover per size defines dimension

$$\text{Dimension} = \lim_{D \to 0} \frac{\text{Log}(N_D^{\text{boxes}})}{\text{Log}(D)}$$

Box Counting

3 Boxes with distance=0.5 cover all points

Number of boxes to cover per size defines dimension

$$\text{Dimension} = \lim_{D \to 0} \frac{\text{Log}(N_D^{\text{boxes}})}{\text{Log}(D)}$$

Box Counting

D=0.25

8 Boxes with distance=0.25 cover all points

Number of boxes to cover per size defines dimension

$$\text{Dimension} = \lim_{D \to 0} \frac{\text{Log}(N_D^{\text{boxes}})}{\text{Log}(D)}$$

Box Counting

Slope 1
(aka D=1)

Number of boxes to cover per size defines dimension

$$\text{Dimension} = \frac{\text{Log}(N_D^{\text{boxes}})}{\text{Log}(D)} \quad (D \rightarrow 0)$$

Box Counting

Our 4 particle jet is fractal like

Number of boxes to cover per size defines dimension

What about a real jet?

Roughly linear scalar separation between quarks and gluons

arxiv/sometime soon

What about a real jet?

Effects from the heavy ion medium can modify this region

arxiv/sometime soon

In Our Detector

Over the past week you have seen the individual components

What if we start to link things together?

We have to account for the other detectors

Electronics feeding calorimeters

Big magnetic

Particle Identification Basics

Credit: Particle Data Group (LBNL)

Particle Flow

Link everything together

Make Particles

We get 5 classes of particles

Are we missing any?

Ordering is key

Particle #1 (Id) & remove (expected)deposit

1. Muons
2. Electrons
3. Charged Hadrons*
4. Photons
5. Neutral hadrons

3 overlapping particle event

Resulting Componnets

- Std::vector<ParticleFlowCandidates>

- ParticleFlowCandate :
  - Particle id
  - Mass
  - p_T
  - Eta
  - Phi

5 types

ATLAS does not (yet) link charged particles w/calo

A lesser known point

- Embedded in particle flow are nuclear interaction

Particle flow looks for displaced secondary vertices and merges tracks into 1 particle

Any complication for PF in ATLAS?

ATLAS

CMS

Any complication for PF in ATLAS?

Magnet position

Need to correct for the energy loss

ATLAS CMS

Using the particles

What is a jet?

A composition of many particles originating from a quark or gluon

What is a jet?

Inside a detector
Tracks in Silicon
From a single vertex
Ecal
Hcal

Jet Reconstruction

- Itrerate over two

Take smallest

$$\Delta R \min(p_T^1, p_T^2)^\alpha$$

α=1 kT
α=0 Cambridge Aachen
α=-1 Anti-kT

Jet Reconstruction

- Itrerate over two

Start small

Take smallest

$$\Delta R \min(p_T^1, p_T^2)^\alpha$$

$\alpha=1$ kT

$\alpha=0$ Cambridge Aachen

$\alpha=-1$ Anti-kT

Jet Reconstruction

- Itrerate over two

Start Close

Take smallest

$$\Delta R \min(p_T^1, p_T^2)^\alpha$$

$\alpha=1$ kT

$\alpha=0$ Cambridge Aachen

$\alpha=-1$ Anti-kT

Jet Reconstruction

- Itrerate over two

Start Big

Take smallest

$$\Delta R \min(p_T^1, p_T^2)^\alpha$$

$\alpha=1$ kT

$\alpha=0$ Cambridge Aachen

$\alpha=-1$ Anti-kT

Jet Reconstruction

- Iterate over two

Take smallest

$$\Delta R \min(p_T^1, p_T^2)^\alpha$$

$\alpha=1$ kT

$\alpha=0$ Cambridge Aachen

$\alpha=-1$ Anti-kT

Now merge initial into a particle

Jet Reconstruction

- Iterate over two

Zooming out

Take smallest

$$\Delta R \min(p_T^1, p_T^2)^\alpha$$

$\alpha=1$ kT

$\alpha=0$ Cambridge Aachen

$\alpha=-1$ Anti-kT

Jet Reconstruction

- Iterate over two

Merge next set

Take smallest

$$\Delta R \min(p_T^1, p_T^2)^\alpha$$

$\alpha=1$ kT

$\alpha=0$ Cambridge Aachen

$\alpha=-1$ Anti-kT

Jet Reconstruction

- Iterate over two

Merge next set

Take smallest

$$\Delta R \min(p_T^1, p_T^2)^\alpha$$

$\alpha=1$ kT

$\alpha=0$ Cambridge Aachen

$\alpha=-1$ Anti-kT

Jet Reconstruction

- Iterate over two

Merge next set

Take smallest

$$\Delta R \min(p_T^1, p_T^2)^\alpha$$

$\alpha=1$ kT

$\alpha=0$ Cambridge Aachen

$\alpha=-1$ Anti-kT

Jet Reconstruction

- Iterate over two

If distance > X (stop)
X=0.4,0.8,...

Take smallest

$$\Delta R \min(p_T^1, p_T^2)^\alpha$$

α=1 kT
α=0 Cambridge Aachen
α=-1 Anti-kT

Jet Reconstruction

- Iterate over two

Done

Take smallest

$$\Delta R \min(p_T^1, p_T^2)^\alpha$$

$\alpha=1$ kT

$\alpha=0$ Cambridge Aachen

$\alpha=-1$ Anti-kT

Why do we use these algorithms?

$$\Delta R \min(p_T^1, p_T^2)^\alpha$$

Need to be able to calculate these with QCD

Collinear safety

Can Randomly happen? When $\Delta R \rightarrow 0$

$\Delta R \min(p_T^1, p_T^2)^\alpha \rightarrow 0$

When $\Delta R \rightarrow 0$

*Wesley Smith

Why do we use these algorithms?

$$\Delta R \min(p_T^1, p_T^2)^\alpha$$

Need to be able to calculate these with QCD

Infrared safety

Can Randomly happen? When $E \to 0$

$\Delta R \min(p_T^1, p_T^2)^\alpha \to 0 \ (p_T \to 0)$

For $\alpha=0$ gluon gets combined with nearest particle $p_T^i \to p_T^i + E(\to 0) = p_T^i$

*Wesley Smith

Why do we use these algorithms?

To calculate anything with a jet we need to observe :

Infrared safety : invariance with random particle w/E→0
Collider safety : invariance with random split ΔR→0

This applies to jet substructure observables too!
…..Well maybe
can you think of a example that breaks IR safety?

Recap

What is a Jet?

Inside a detector

Charged Pileup can be separated by vertexing

Neutral Pileup cannot*

*Fast timing/Depth reconstruction can help

Detector
Detector Surface
ϕ
6
4
2
0
-4
-2
0
2
4
η
10^2
10
1
10^{-1}
pT (GeV)
>200 GeV
50-200 GeV
25-50 GeV

We also have pileup

We also have pileup

- Filtering the interesting info

We also have pileup

- Filtering the interesting info

Pileup

- While Pileup in pp has some differences
  Many of the properties of pileup are similar to UE fluctuations

>200 GeV

50-200 GeV

25-50 GeV

Jet Energy Correction

- Correcting to truth

How do we shape our jet against the UE?
Why did CMS switch to AK4?

Part 1
Hadronic Resonances

Part 1
Hadronic Resonances

III.4.2. $H \to b\bar{b}$

The production mode $gg \to H$ with $H \to b\bar{b}$ is overwhelmed by the background from the inclusive production of $p\bar{p} \to b\bar{b} + X$ via the strong interaction. The associated

Particle Data Group 2016
(High Energy Physics Handbook)

History

- With the first generation of substructure
  - We were able to resolve overlapping objects
  - In regions where clear signatures were expected
    - Found them and characterized the

- Led to a
  - Wealth of new physics bounds
  - Transformative in our approach to LHC

A few Questions remained uncertain

- If you look closely
  - Jesse is setting up the beginning of a problem
  - The same problem that PDG said was impossible

In the rest of this talk
Explain how we
solved this problem

Standing on the
shoulders of giants

EVENT DISPLAYS!

This is actually what it looks like in the detector

p_T = 466 GeV
double-b = 0.95
m_{SD} = 126.2 GeV
$N_2^{1\ DDT}$ = -0.07

What you know now:

Jets are signatures of quarks and gluons

Lets go back and try again?

To see if its useful lets target a core process:

A W boson decaying to quarks

What is a hadronic resonance?

A Heavy Object that decays to quarks:

A simple example is a φ to Kaons

Kaons decay to 3 Pions

What is a hadronic resonance?

A Heavy Object that decays to quarks:

a complicated example is a W boson to quarks

High energy quarks decay to a shower of paricles

What makes this hard?

What is a hadronic resonance?

A Heavy Object that decays to quarks:

a complicated example is a W boson to quarks

High energy quarks decay to a shower of paricles

What makes this hard?

What is a hadronic resonance?

A Heavy Object that decays to quarks:

a complicated example is a W boson to quarks

High energy quarks decay to a shower of paricles

What makes this hard?

What makes it Hard?

Two main things:

Trying to select Heavy Ws to quarks is hard
Need a way to remove the background

What makes it Hard?

Two main things:

Overlap

$$e_3^{(\beta=2)} = \sum_{i<j<k\in\text{jet}} z_i z_j z_k \theta_{ij}^2 \theta_{ik}^2 \theta_{jk}^2 \approx \sum_{i<j} z_i z_j \theta_{ij}^2 \theta_i^2 \theta_j^2$$

$$\approx \sum_{i<j} z_i z_j \max(\theta_i^2, \theta_j^2) \theta_i^2 \theta_j^2 \approx \sum_{i<j} \rho_i \rho_j \max(\theta_i^2, \theta_j^2),$$

To resolve the overlap we resort to aggregate combinations of particles

What you know now:

Jets of the Standard Model

What you know now:

Jets of the Standard Model

High boost

Let's consider a W-decay

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$$

At ~rest

$\Delta R(q\bar{q}) \approx \pi$

low W-momentum (p_T~100 GeV)

$\Delta R(q\bar{q}) \approx \pi/2$

high boost (p_T>200 GeV)

$\Delta R(q\bar{q}) < 0.8$

$H \to b\bar{b}$
b-jet
$Z \to e^+e^-$
e
b-jet
e

CMS Experiment at the LHC, CERN
Data recorded: 2017-Oct-20 03:55:39.135168 GMT
Run / Event / LS: 305313 / 624767783 / 361
$H \to b\bar{b}$
b-jet
b-jet
q-jet or g-jet

CMS Experiment at the LHC, CERN
Data recorded: 2017-Oct-20 03:55:39.135168 GMT
Run / Event / LS: 305313 / 624767783 / 361
$H \to b\bar{b}$
q-jet or g-jet
b-jet
b-jet
Higgs
b
$\bar{b}$
m_{jet} ~ 125 GeV
E_{jet} ~ 500 GeV
2-prong jet
b-decay inside

Jet substructure

Resolving emissions within the jet

Jet grooming

Jet tagging

m_{jet} ~ 125 GeV
E_{jet} ~ 500 GeV
2-prong jet
b-decay inside

Jet grooming

Removing a jet’s soft and wide-angle radiation

Grooming declustering

Original jet

=

In a clustering tree

Grooming declustering

Groomed jet

=

Groomed clustering tree

- Grooming algorithms iteratively decluster a jet
- The soft-drop algorithm removes lowest p_T sub-jets failing pairwise condition

$$z > z_{cut}\left(\frac{\Delta R_{12}}{R}\right)^{\beta} \quad \text{with} \quad z = \frac{\min(p_T^1, p_T^2)}{p_T^1 + p_T^2}$$

Calculating groomed mass

- Improves the mass *resolution* of a jet (removes away excess radiation)
- Brings the mass of background jets (QCD) to low mass

Jet tagging

Characterizing a jet’s internal decay

Signatures:

quark jet
u, d, s

gluon-jet

X-jet

W-jet

Z-jet

Higgs jet

Top jet

N-subjetiness

Counting the number of "prongs" for almost a decade

Energy weight

$$\tau_N(\mathcal{J}) = \min_{N \text{ axes}} \sum_i E_i \min\{\theta_{1,i}, \theta_{2,i}, \ldots, \theta_{N,i}\}$$

Angular weight w. respect to one axis

Sum over all particles

How likely to be consistent w. 3 axes?

How likely to be consistent w. 2 axes?

N-subjetiness

Counting the number of "prongs" for almost a decade

$$\tau_N(\mathcal{J}) = \min_{N \text{ axes}} \sum_i E_i \min\{\theta_{1,i}, \theta_{2,i}, \ldots, \theta_{N,i}\}$$

Whats the deal with deep learning?

There are 30-200 entangled particles In a jet

$$\mathcal{C}_N = \sum_{k=0}^{N} \binom{N}{k} \sum_{i_1=0}^{1} \cdots \sum_{i_k=0}^{1} \lambda^{\sum_{a=1}^{k} i_a} (1-\lambda)^{k-\sum_{a=1}^{k} i_a} E_1^k$$
$$\times \sum_{i_{k+1}=2}^{M} \cdots \sum_{i_N=2}^{M} C_{i_1 \cdots i_N}(\hat{p}_{i_1}^{\mu}, \ldots, \hat{p}_{i_N}^{\mu}) \prod_{j=k+1}^{N} E_{i_j}$$
$$= \sum_{k=0}^{N} \binom{N}{k} \sum_{\ell=0}^{k} \binom{k}{\ell} \lambda^{\ell} (1-\lambda)^{k-\ell} E_1^k$$
$$\times \sum_{i_{k+1}=2}^{M} \cdots \sum_{i_N=2}^{M} C_{\underbrace{0\cdots0}_{\ell}\underbrace{1\cdots1}_{k-\ell} i_{k+1}\cdots i_n}(\hat{p}_1^{\mu}, \ldots, \hat{p}_1^{\mu}, \hat{p}_{i_{k+1}}^{\mu}, \ldots, \hat{p}$$

Deep Learning has made it possible to decode some of the finer features of a jet

Deep Learning taggers

The rise of machine learning: two examples in CMS

Particles can be sent to NN

Ignores multi-particle correlation

Represent jets as images

Images are not Lorentz invariant

How to better represent a jet in ML?

How its evolved over time

- Deep learning for substructure is rapidly growing field

2016

Images
(not lorentz invariant)

2018

Particles and SVs
with 4-vectors+features

Particles
(limited correlations)

2020

Graphs
(Particles+correlations)

Current collaboration results

Progressively moving towards use of more info

J.Krupa, D. Rankin, S. Siegel (UROP),K. Llewellen

Why are we so excited?

A collision event

A collection of particles in momentum space

Jet as a point cloud

A collection of points in position space (x,y,z)

Particle-Net

A graph neural network using particle-clouds

- For a jet: we have 2D information of particles, but also energy/momenta, charge/particle ID, track quality/impact parameter, etc.

Jet as a particle cloud Each cloud is a graph Perform convolution Get ID

Unordered representation

Edges: connections between point and k-NN points

Vertices: point (Particle or secondary vertex)

EdgeConv outputs a set of learned features for each point

Binary classification from all learned features

Particle-Net

A graph neural network using particle-clouds

- A step in the right direction: Networks and correlations can be included (the most powerful discrimination in CMS)
- Still a long way to go:
  - More on DL by Dylan

Lets go back and try again?

To see if its useful lets target a core process:

A W boson decaying to quarks

In an experiment

Find:

A W boson decaying to quarks
A low p_T we resolve them as jets

In an experiment

Find:

Resolved as a tiny bump on a huge background

A W boson decaying to quarks

A low p_T we resolve them as jets

Was last observed 30 years

Result was not very convincing

Have not seen this since

1988

How do we reduce background:

Resolved as a tiny bump on a huge background

Whats the first goal?

Require W to be a high p_T

Isolating the W

How do we reduce background:

The core issue is that the background is too large
Quark and gluon interactions are large!

Isolating the W

The core issue is that the background is too large

Whats the first goal?

Require W to be a high p_T

Quarks decaying from high p_T W nearby each other

Resolve this as a single jet

Recoiling jet off W gives the W high p_T

Whats the first goal?

Require W to be a high p_T

Have not observed a single W jet boson event from SM produce
(Have only seen W jets in top events)

Performing a basic selection

Tools developed with the first rounds of jet substructure did not lead to a significant signature

Hard to convince anybody that bump is at 80
With another big bump at a higher mass

Part 2
Decoding Data

A Game

Lets play a game

- Lets try to inject a signal likes this

Into a background 50 times larger than this
This signal looks very much like a W peak

Find the signal

- Can you see the signal?

Find the signal

- Can you see the signal?

Here is the signal at 125 GeV on a flat background

That was an easy one!

Find the signal

- Can you see the signal?

Here is the signal at 125 GeV on a flat background

Now the background peaks just after the signal

Find the signal

- Can you see the signal?

Here it is. This one is a bit harder.

Note exact same amount of background and signal

How about now?

- Can you see the signal?

How about now?

- Can you see the signal?

Background

Note exact same amount of background and signal

Whats best background?

All variables yield roughly the same performance

Exactly the same background yield

Exactly the same signal yield

What if we could tune background shape to what we want?

What would be our best choice of shape?

Whats best background?

All variables yield roughly the same performance

Exactly the same background yield

Exactly the same signal yield

What if we could tune background shape to what we want?

What would be our best choice of shape?

Consider trying to isolate a signal

Extracting a Signal

Signal Events

Background

Select Events

And Fit Mass

Not Good Enough!

Extracting a Signal

- Signal Events
- Background

Select Events

Cut tighter!

And Fit Mass

Use Deep Learning

Signal Events

Background

Select Events

And Fit Mass

Still Not Enough!

Use Deep Learning

- Signal Events
- Background

Select Events

Can't separat sig+bkg

And Fit Mass

Conceptually what was done?

Now its good!

Conceptually what was done?

Mass

A’

Mass

Now its good!

Jet Mass distribution

What makes this complicated?

$$A'(in_1, in_2, \ldots, in_{1000})$$

The actual variable we cut on can be a function of 1000 different variables

No Joke
(I will get back to that)

Mass

What makes this complicated?

To start with we relied on a standard:

$A'(in_1, in_2, \ldots, in_{1000}) = \tau_2/\tau_1$

+Brute force decorrelation

Mass

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min\{\Delta R_{1,k}, \Delta R_{2,k}, \cdots, \Delta R_{N,k}\}$$

What makes this complicated?

After Jesse's paper we switched to:

$$A'(in_1, in_2, \ldots, in_{1000}) = \frac{{}_2e_3^{(\beta)}}{\left(e_2^{(\beta)}\right)^2}(N_2)$$

+Brute force decorrelation

Mass

$$e_3^{(\beta=2)} = \sum_{i<j<k\in\text{jet}} z_i z_j z_k \theta_{ij}^2 \theta_{ik}^2 \theta_{jk}^2 \approx \sum_{i<j} z_i z_j \theta_{ij}^2 \theta_i^2 \theta_j^2$$

$$\approx \sum_{i<j} z_i z_j \max(\theta_i^2, \theta_j^2) \theta_i^2 \theta_j^2 \approx \sum_{i<j} \rho_i \rho_j \max(\theta_i^2, \theta_j^2),$$

J.Krupa

Our current result

http://cms-results.web.cern.ch/cms-results/public-results/publications/EXO-18-012/

Only through above strategies could we get here

Ultimately it meant we could safely cut tigher

J.Krupa

Our Current Result

Starting to get real precision

Physics Implications

advantage of the harder transverse momentum spectrum of the $gg \to H$ production mode with respect to the QCD background, a search for high p_T Higgs boson decaying to a pair of b quarks in association with an energetic Initial State Radiation (ISR) jet, has been performed by ATLAS [155] and CMS [156]. For this analysis with the Run 2 data, ATLAS and CMS require jets clustered with

Particle Data Group 2020
(High Energy Physics Handbook)

Higgs at High Energy

Process with the most events is the best at probing highest energy Higgs bosons

Can now measure Higgs process in the blue region

Could not before

Note: The prediction in this region is not well known

State of the Universe

- What does this have to do with it?

classical Higgs potential

stable

unstable/metastable

High energy regime *potentially* tells us about the stability of the Higgs potential

D. Rankin, S.E. Park

Higgs at high p_T

Relied on the Z to perform Calibration $\mu_Z=1.01^{+0.24}_{-0.20}$

μ is wrt to LHC Higgs WG

Large Excess at Higgs mass $\mu_H=3.7^{+1.6}_{-1.5}$ (2.5σ)

D. Rankin, S.E. Park

Delving in more detail

Trend toward higher Energy

Consistent with modified coupling

This regime is difficult to calculate

Large excess shows an interesting trend

What does this mean?

Large excess with a trend means…..

Our prediction for high energy Higgs bosons is off
Higgs boson production at high energy is HARD

There is the potential of some new physics

We are not yet at the point where we can make a significant claim

Low mass (qq) resonances

Looking for a resonance

At the LHC

$p_T < 200$ GeV

$m_X > 500$ GeV

If m_X < 500 GeV, event does not get saved

LHC has tight energy thresholds (mass thresholds)

p_T < 200 GeV

m_X < 500 GeV

How to save these events

Boost them!

$p_T > 450$ GeV

$m_X < 500$ GeV

What the process looks like

And what you'll be looking for

p_T > 450 GeV

m_W ~ 80.4 GeV

Your goal: reproduce LHC results (or improve them)

With 8 TeV data

@ 8 TeV (2015)

Finding W

Finding Z

Your goal: reproduce LHC results (or improve them)

With 8 TeV data

@ 8 TeV (2015)

Finding W

Finding Z

Project 2

Finding W(qq) and Z(qq) in 8 TeV LHC data

Introduction

In this lab we will investigate W bosons produced in the LHC's 8 TeV proton proton collisions. These samples were produced 4 years ago in a fun experiment that opened up the option of performing low mass resonance searches at the LHC. The studies done then have led to a wealth of results from both LHC experiments, ATLAS and CMS. To understand how this study works, we first need to introduce a few concepts.

What is a measurement?

- Lets say we have a process that we believe behaves like this
  - This process is what we call our **Prior**

What is a measurement?

- We have assumed a model for this process
  - Example what is the direction of your eye reading slide?

Given our Prior we say that this measurement has a z-score of 2

What happens with many measurements?

- What if the distribution we observe is different?

What happens with many measurements?

- What if the distribution we observe is different?

What happens with many measurements?

- What if the distribution we observe is different?

Are we missing something?

In reality we could be inundated with Data and we only have a biased selection

Are we missing something?

Monty Hall Problem

There is a car behind one of these doors?
Which one do you choose?

Monty Hall Problem

There is a car behind one of these doors?
Which one do you choose?

Monty Hall Problem

Now we open one door and it doesn’t have a car
Do you stay on Door 1 or do you go to Door 3

Monty Hall Problem

Now we open one door and it doesn’t have a car
Do you stay on Door 1 or do you go to Door 2

Monty Hall Problem

Now we open one door and it doesn’t have a car
Do you stay on Door 1 or do you go to Door 2

Monty Hall Problem

- Let’s make a deal
- https://www.youtube.com/watch?v=iBdjqtR2iK4
- https://medium.com/@NickDoesData/applying-bayes-theorem-simulating-the-monty-hall-problem-with-python-5054976d1fb5

Monty Hall Problem

What happens with many measurements?

- How did we go from Prior to Posterior?
- For many cases this is an application of bayes theorem

$$P(\mathcal{H}|\mathcal{D}) = \frac{P(\mathcal{D}|\mathcal{H})\,P(\mathcal{H})}{P(\mathcal{D})}$$

$$P(\text{hypothesis}|\text{data}) = \frac{P(\text{data}|\text{hypothesis})\,P(\text{hypothesis})}{P(\text{data})}$$

Prior

Likelihood

Posterior

-

What happens with many measurements?

- How did we go from Prior to Posterior?
- For many cases this is an application of bayes theorem

$$P(\mathcal{H}|\mathcal{D}) = \frac{P(\mathcal{D}|\mathcal{H})\,P(\mathcal{H})}{P(\mathcal{D})}$$

$$P(\text{hypothesis}|\text{data}) = \frac{P(\text{data}|\text{hypothesis})\,P(\text{hypothesis})}{P(\text{data})}$$

Prior

Likelihood

Posterior

Normalizer

$$\text{posterior} \propto \text{likelihood} \times \text{prior}$$

What happens with many measurements?

- How did we go from Prior to Posterior?
- For many cases this is an application of bayes theorem

$$P(\mathcal{H}|\mathcal{D}) = \frac{P(\mathcal{D}|\mathcal{H})\boxed{P(\mathcal{H})}}{P(\mathcal{D})}$$

Prior

$$P(\text{hypothesis}|\text{data}) = \frac{P(\text{data}|\text{hypothesis})\boxed{P(\text{hypothesis})}}{P(\text{data})}$$

Posterior

- Lets do a real world example
  - LHC Excess

A Strange LHC Excess

Cornell University

the Simo

arXiv > hep-ph > arXiv:2203.13180

Search...

Help | Advan

High Energy Physics – Phenomenology

[Submitted on 24 Mar 2022 (v1), last revised 21 Aug 2022 (this version, v2)]

Mounting evidence for a 95 GeV Higgs boson

Thomas Biekötter, Sven Heinemeyer, Georg Weiglein

In 2018 CMS reported an excess in the light Higgs-boson search in the diphoton decay mode at about 95GeV based on Run 1 and first year Run 2 data. The combined local significance of the excess was $2.8\,\sigma$. The excess is compatible with the limits obtained in the ATLAS searches from the diphoton search channel. Recently, CMS reported another local excess with a significance of $3.1\,\sigma$ in the light Higgs-boson search in the di-tau final state, which is compatible with the interpretation of a Higgs boson with a mass of about 95GeV. We show that the observed results can be interpreted as manifestations of a Higgs boson in the Two-Higgs Doublet Model with an additional real singlet (N2HDM). We find that the lightest Higgs boson of the N2HDM can fit both excesses simultaneously, while the second-lightest state is such that it satisfies the Higgs-boson measurements at 125GeV, and the full Higgs-boson sector is compatible with all Higgs exclusion bounds from the searches at LEP, the Tevatron and the LHC as well as with other theoretical and experimental constraints. Finally, we demonstrate that it is furthermore possible to accommodate the excesses observed by CMS in the two search channels together with a local $2.3\,\sigma$ excess in the $b\bar{b}$ final state observed at LEP in the same mass range.

(a) Class 0

(b) Class 1

A Strange LHC Excess

Cornell University

the Simo

arXiv > hep-ph > arXiv:2203.13180

Search...

Help | Advan

High Energy Physics – Phenomenology

[Submitted on 24 Mar 2022 (v1), last revised 21 Aug 2022 (this version, v2)]

Mounting evidence for a 95 GeV Higgs boson

Thomas Biekötter, Sven Heinemeyer, Georg Weiglein

In 2018 CMS reported an excess in the light Higgs-boson search in the diphoton decay mode at about 95GeV based on Run 1 and first year Run 2 data. The combined local significance of the excess was $2.8\,\sigma$. The excess is compatible with the limits obtained in the ATLAS searches from the diphoton search channel. Recently, CMS reported another local excess with a significance of $3.1\,\sigma$ in the light Higgs-boson search in the di-tau final state, which is compatible with the interpretation of a Higgs boson with a mass of about 95GeV. We show that the observed results can be interpreted as manifestations of a Higgs boson in the Two-Higgs Doublet Model with an additional real singlet (N2HDM). We find that the lightest Higgs boson of the N2HDM can fit both excesses simultaneously, while the second-lightest state is such that it satisfies the Higgs-boson measurements at 125GeV, and the full Higgs-boson sector is compatible with all Higgs exclusion bounds from the searches at LEP, the Tevatron and the LHC as well as with other theoretical and experimental constraints. Finally, we demonstrate that it is furthermore possible to accommodate the excesses observed by CMS in the two search channels together with a local $2.3\,\sigma$ excess in the $b\bar{b}$ final state observed at LEP in the same mass range.

(a) Class 0

(b) Class 1

How to View it

Bayesian vs Frequentists

Probability (mathematics)

$$P(H|D) = \frac{P(D|H)P(H)}{P(D)}$$

Everyone uses Bayes' formula when the prior $P(H)$ is known.

Bayesian path

Frequentist path

Statistics (art)

$$P_{\text{Posterior}}(H|D) = \frac{P(D|H)P_{\text{prior}}(H)}{P(D)}$$

Bayesians require a prior, so they develop one from the best information they have.

Likelihood $L(H;D) = P(D|H)$

Without a known prior frequentists draw inferences from just the likelihood function.

Bayesian View

My posterior is the result of some selection bias of a deeper model
I just need to fix my conditional probabilities

Frequentist View

All we know is our result gives us a gaussian at 2

Bayesian stats people are a bunch of

<u>smart asses guessing a prior</u>

Bayesian Mean=0 and Std Dev=1, Frequentist Mean=2.002+/-0.06,StdDev=1.02+/-0.03

What if models is incomplete?

- https://www.youtube.com/watch?v=Ku42Iszh9KM
- Sometimes there are other parameters not included in our model

What if models is incomplete?

- https://www.youtube.com/watch?v=Ku42Iszh9KM
- Sometimes there are other parameters not included in our model

Bayesian view:
Each parameter adds a dimension we probe

Frequestist View:
Make sure your model is complete

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right).$$

Float this parameter

Nuisance

External Nuisances

$$-2\log\left(\mathcal{L}(x|\vec{\theta})\right) = -2\log\left(\mathcal{L}(x|\vec{\theta})\right)_{\text{original}} + \frac{\left(\sin\theta_{23} - \sin\theta_{23}^{\text{best}}\right)^2}{\sigma^2_{\sin\theta_{23}}} + \frac{\left(\Delta m^2_{23} - \Delta m^{2\,\text{best}}_{23}\right)^2}{\sigma^{2\,\text{best}}_{\Delta m^2_{12}}}$$

Nuisance

CMS Supplementary
35.9 fb-1 (13 TeV)
1 tt̄+bb̄ cross section (50%)
2 b tagging: charm (linear)
3 tt̄H cross section (renorm./fact. scales)
4 jet energy scale (1)
5 tt̄+2b cross section (50%)
6 b tagging: lf fraction
7 b tagging: lf stats (quadratic)
8 μR scale (tt̄)
...
−2 −1 0 1 2
(θ̂-θ0)/Δθ
−0.1 0 0.1
Δμ
Pull
+1σ Impact
-1σ Impact

Nuisance

Post-fit impact on μ:
θ = θ̂+Δθ̂
θ = θ̂-Δθ̂
Nuis. Param. Pull
ATLAS
√s = 13 TeV, 36.1 fb-1
tt̄+jets PS and hadronization
tt̄+jets NLO generator
tt̄+≥1c normalization
tt̄+≥1b NLO reweighting
tt̄+≥1b MPI normalization
b-tagging efficiency (NP I)
JES (flavor composition)
tt̄+light /≥1c NNLO top-pT corr.
W/Z+jets normalization (9j)
tt̄+≥1b normalization
–2 –1.5 –1 –0.5 0 0.5 1 1.5 2
(θ̂-θ0)/Δθ

Lecture 10: Hypothesis Testing

Hypothesis Test

- Null Hypothesis :
  - $p_a(x) = p_b(x)$
- Alternative Hypothesis
  - $p_a(x) \neq p_b(x)$
- Answers these questions:
  - **Is my data consistent?**
  - **Can I detect an effect?**
  - **Did I just make a discovery?**

Prior Posterior

$p_a(x)$ $p_b(x)$

Hypothesis Test

Lets Get Back to the original

- How do I tell if my measurements are consistent?

What is a measurement?

- We have assumed a model for this process
  - Example what is the direction of your eye reading slide?

Given our Prior we say that this measurement has a z-score of 2

Think Like a Frequentist

- How do I tell if my measurements are consistent?

Frequentist: chi2-test

- To test if two histograms are consistent
  - $$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} = N \sum_{i=1}^{n} \frac{(O_i/N - p_i)^2}{p_i}$$

where

χ^2 = Pearson's cumulative test statistic, which asymptotically approaches a χ^2 distribution.

O_i = the number of observations of type *i*.

N = total number of observations

$E_i = Np_i$ = the expected (theoretical) count of type *i*, asserted by the null hypothesis that the fraction of type *i* in the population is p_i

n = the number of cells in the table.

If we have a continuous distribution we can turn it into a histogram with zero uncertainty, there are a lot of ways to do this

chi2 consistency

- Variation as a function of sample size
  - Compute chi2 consistency between histogram and observed
  - Larger p-value with larger n-samples

KS-test

- Komolgorov-Smirnov Test
  - Estimate the maximum difference between observed and predicted cumulative distribution functions and compare with expectations.
  - An unbinned way to check histograms are equal

$$\widehat{F}_n(t) = \frac{\text{number of elements in the sample} \leq t}{n} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{X_i \leq t},$$

$$D_{n,m} = \sup_x |F_{1,n}(x) - F_{2,m}(x)|,$$

Supremum Function (ie max difference)