FACULTY OF SCIENCES

journal club 20/03/2024

What Makes A Discovery?

Dan Hooper

interpreted by Jules Vandenbroeck *(or at least I tried to)*

GHENT UNIVERSITY

FACULTY
OF SCIENCES

5-σ significance

Thank you for listening

***interpreted by Jules Vandenbroeck** (or at least I tried to)*

GHENT
UNIVERSITY

situate the "paper": what makes a discovery?

proceedings by **Dan Hooper**, cosmologist and particle physicist

He attended a Nobel symposium on the topic of dark matter and gave a talk on the status of the Galactic Center Gamma-Ray Excess

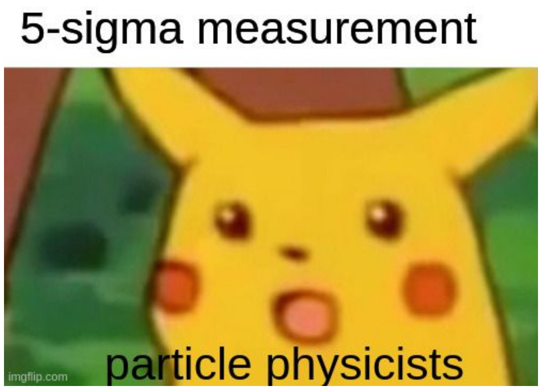
Instead of writing about his talk in the proceedings, he wrote on another series of discussions that took place at that meeting.

> *"What would have to occur before we, as a scientific community, would be willing to declare that we had discovered the nature of dark matter?"*

disclaimer: this presentation is partially philosophical (opinions matter) and you can interrupt me if you don't agree with my or Dan's interpretation! Limited conclusions are presented and all results are up for the interpretation of the listener.

What Makes a Discovery a Discovery?

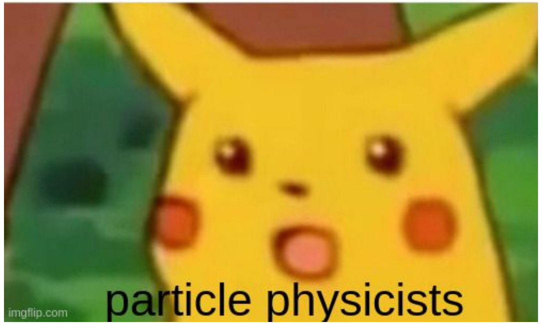
"In the field of particle physics, the word discovery is sometimes treated as a synonym for one or more measurements that are discrepant with the predictions of the null hypothesis at a level of at least 5 standard deviations."



What Makes a Discovery a Discovery?

“In the field of particle physics, the word discovery is sometimes treated as a synonym for one or more measurements that are discrepant with the predictions of the null hypothesis at a level of at least 5 standard deviations.”

5-sigma measurement



Does this definition hold up in practice? (and should it?)

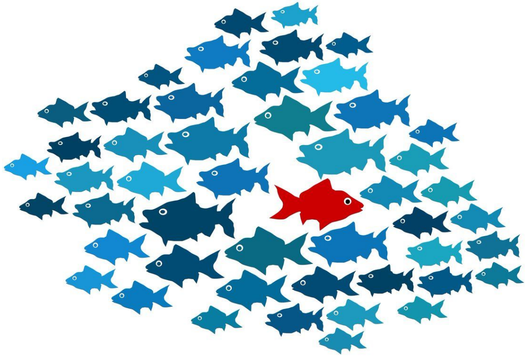
questions to answer

- are all discoveries covered by the same level of scrutiny?
- What goes into the response of the scientific community to new data and experimental anomalies?

What Makes a Discovery a Discovery?

How can an anomaly appear in the data and lead to a discovery?

1) statistical fluctuation
2) underestimation of systematic uncertainties
3) error in the measurement
4) underestimation of trials factor (“look elsewhere” effect)
5) new physics!

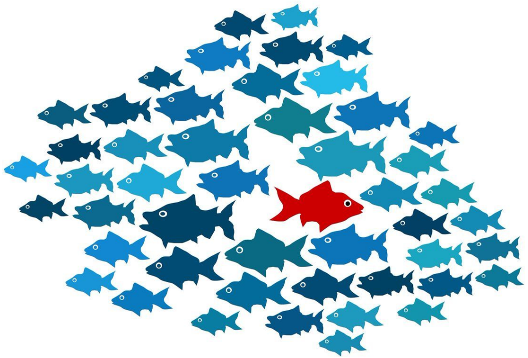


GHENT UNIVERSITY

What Makes a Discovery a Discovery?

How can an anomaly appear in the data and lead to a discovery?

1) statistical fluctuation
2) underestimation of systematic uncertainties
3) error in the measurement
4) underestimation of trials factor ("look elsewhere" effect)
5) new physics!



discussion

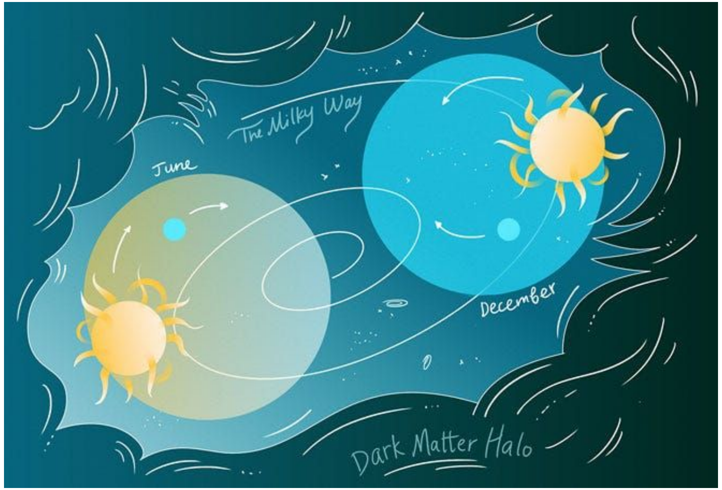
Is every 5-σ significance anomaly considered a discovery?

eg.: DAMA Collaboration, significance of 8.9σ[1]

→does the community consider this experiment to have detected particles of dark matter?

GHENT UNIVERSITY

intermezzo 1: the DAMA Collaboration



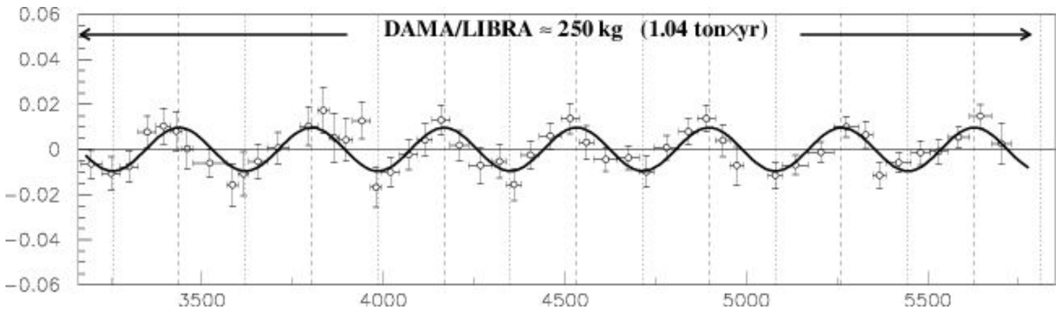
Direct dark matter detector

Goal: detect dark matter particles in the galactic halo with scintillation detectors

Signature: annual revolution of event rate in DAMA detector due to our motion around the Sun

Result: annual modulation of the event rate with a significance of 8.9σ.

Controversy: signal lies in a parameter range excluded by the results of other experiments

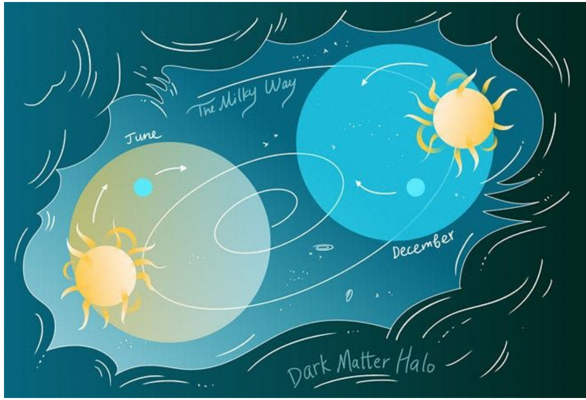


GHENT UNIVERSITY

What Makes a Discovery a Discovery?

How can an anomaly appear in the data and lead to a discovery?

1) statistical fluctuation
2) underestimation of systematic uncertainties
3) error in the measurement
4) underestimation of trials factor (“look elsewhere” effect)
5) new physics!



discussion

Is every 5-σ significance anomaly considered a discovery?

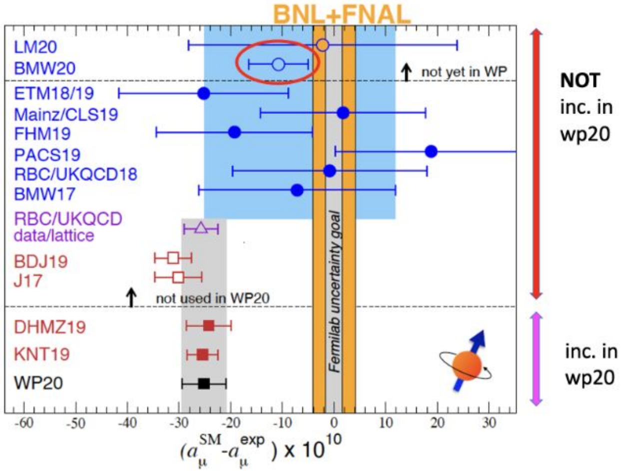
eg.: DAMA Collaboration, significance of 8.9σ[1]
→does the community consider this experiment to have detected particles of dark matter?

GHENT UNIVERSITY

What Makes a Discovery a Discovery?

How can an anomaly appear in the data and lead to a discovery?

1) statistical fluctuation
2) underestimation of systematic uncertainties
3) error in the measurement
4) underestimation of trials factor (“look elsewhere” effect)
5) new physics!



discussion

Is every 5-σ significance anomaly considered a discovery?

eg.: Muon’s anomalous magnetic moment tension with theory of 5.1σ[2]
→does the community consider this experiment to have discovered an anomalous magnetic moment? (at the time)

GHENT UNIVERSITY

Trials and Trials of Trials

"look elsewhere" effect

"When searching for a new resonance somewhere in a possible mass range, the significance of observing a local excess of events must take into account the probability of observing such an excess anywhere in the range"[3]

Trials and Trials of Trials

"look elsewhere" effect

"When searching for a new resonance somewhere in a possible mass range, the significance of observing a local excess of events must take into account the probability of observing such an excess anywhere in the range"[3]

Excess of diphoton event with inv. mass of ~750 GeV at the LHC?

- Initially, the (local) significance of this 750 GeV excess was reported to be 3.9σ (3.4σ) by ATLAS (CMS)
  → 0.005% (0.03%) probability to obtain this signal *(must be true?)*

- Taking into account for the "look elsewhere" effect (search in invariant mass range) significance went to 2.1σ (1.6σ)
  → 1.79% (5.48%) probability to obtain a signal at any mass point

GHENT UNIVERSITY

Trials and Trials of Trials

"look elsewhere" effect

"When searching for a new resonance somewhere in a possible mass range, the significance of observing a local excess of events must take into account the probability of observing such an excess anywhere in the range"[3]

Excess of diphoton event with inv. mass of ~750 GeV at the LHC?

- Initially, the (local) significance of this 750 GeV excess was reported to be 3.9σ (3.4σ) by ATLAS (CMS)
  → 0.005% (0.03%) probability to obtain this signal *(must be true?)*

- Taking into account for the "look elsewhere" effect (search in invariant mass range) significance went to 2.1σ (1.6σ)
  → 1.79% (5.48%) probability to obtain a signal at any mass point

"Treating the ATLAS observation as a confirmation of the CMS excess"

→ *"look elsewhere" effect only for CMS probability of obtaining this result with SM:*
0.03% x 1.79% ~ 5 x 10^{-6}
4.6σ!!

Trials and Trials of Trials

"Although this excess generated a great deal of excitement at the time, I don't know of anyone who thought that it was especially likely that this excess was due to new physics. I certainly never heard anyone express the view that the odds of this were anywhere close to $1 - (5 \times 10^{-6}) \approx 99.9995\%$*"*

discussion

(Why) was this the case?

GHENT UNIVERSITY

Trials and Trials of Trials

"Although this excess generated a great deal of excitement at the time, I don't know of anyone who thought that it was especially likely that this excess was due to new physics. I certainly never heard anyone express the view that the odds of this were anywhere close to $1 - (5 \times 10^{-6}) \approx 99.9995\%$"

discussion

(Why) was this the case?

underestimation of trials factor

thousands of searches for new physics at the LHC
→ statistical fluctuations can happen

GHENT UNIVERSITY

Trials and Trials of Trials

"Although this excess generated a great deal of excitement at the time, I don't know of anyone who thought that it was especially likely that this excess was due to new physics. I certainly never heard anyone express the view that the odds of this were anywhere close to $1 - (5 \times 10^{-6}) \approx 99.9995\%$"

discussion

Why was this the case?

underestimation of trials factor

did they expect an excess in this region?

thousands of searches for new physics at the LHC
→ statistical fluctuations can happen

GHENT UNIVERSITY

Theory Bias or Theory Guidance?

In contrast to the Higgs boson discovery

first Higgs-boson results

CMS:	4.9σ
ATLAS:	5.0σ

Almost immediately, most of us in the particle physics community began to refer to this as the "discovery of the Higgs boson." ~ Dan Hoover

GHENT UNIVERSITY

Theory Bias or Theory Guidance?

In contrast to the Higgs boson discovery

first Higgs-boson results

CMS:	4.9σ
ATLAS:	5.0σ

Almost immediately, most of us in the particle physics community began to refer to this as the "discovery of the Higgs boson." ~ D

GHENT UNIVERSITY

Theory Bias or Theory Guidance?

The proceedings provide two main arguments:

(factual)

- Higher statistical significance of the discovery
  → to overcome the look elsewhere penalty

(controversial, but more interesting!)

- the discovery of the Higgs-boson by the LHC was expected by the community

"In other words, most of us evaluated the prior probability of the Higgs boson to be much greater than that of something capable of generating the 750 GeV excess."

GHENT UNIVERSITY

Theory Bias or Theory Guidance?

The proceedings provide two main arguments:

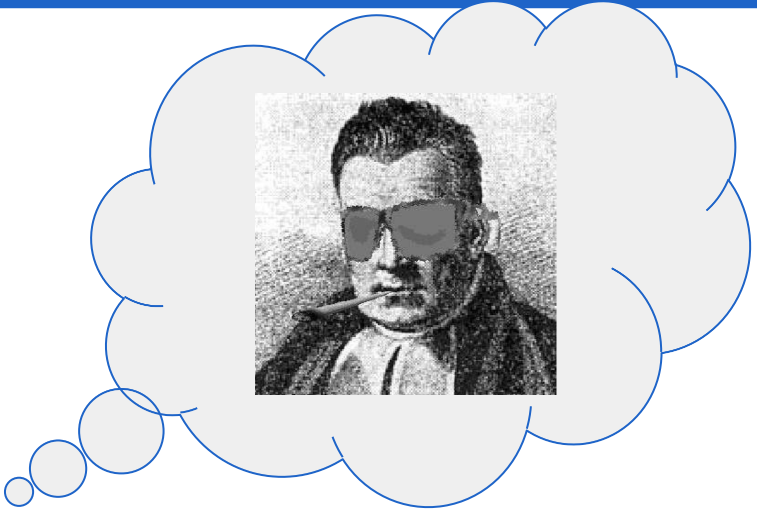
(factual)

- Higher statistical significance of the discovery
  → to overcome the look elsewhere penalty

(controversial, but more interesting!)

- the discovery of the Higgs-boson by the LHC was expected by the community

"In other words, most of us evaluated the prior probability of the Higgs boson to be much greater than that of something capable of generating the 750 GeV excess."



GHENT UNIVERSITY

Theory Bias or Theory Guidance?

The proceedings provide two main arguments:

(factual)

- Higher statistical significance of the discovery
  → to overcome the look elsewhere penalty

(controversial, but more interesting!)

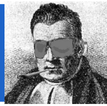
- the discovery of the Higgs-boson by the LHC was expected by the community


thomas bayes

“In other words, most of us evaluated the prior probability of the Higgs boson to be much greater than that of something capable of generating the 750 GeV excess.”

GHENT UNIVERSITY

Intentional Bayesianism

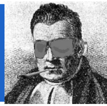


Bayes’ theorem in the pursuit of science

Science is to determine whether a theory is true or false, with prior knowledge or assumptions and newly acquired data (D).

GHENT UNIVERSITY

Intentional Bayesianism



Bayes' theorem in the pursuit of science

Science is to determine whether a theory is true or false, with prior knowledge or assumptions and newly acquired data (D).

likelihood that the dataset D would be obtained if the theory is true.
Previous results can impact this probability

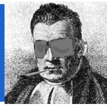
$$P(T_i|D) = \frac{P(D|T_i)\,P(T_i)}{P(D)}.$$

prior (knowledge)
prior probability that the theory is true

goal
(posterior) probability that the theory is true

GHENT UNIVERSITY

Intentional Bayesianism



Bayes' theorem in the pursuit of science

Science is to determine whether a theory is true or false, with prior knowledge or assumptions and newly acquired data (D).

likelihood that the dataset D would be obtained if the theory is true.
Previous results can impact this probability

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

goal

(posterior) probability that the theory is true

prior (knowledge)

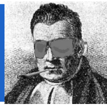
subjective by nature and dependent on a variety of factors:

- theory's motivation
- naturalness
- plausibility
- …

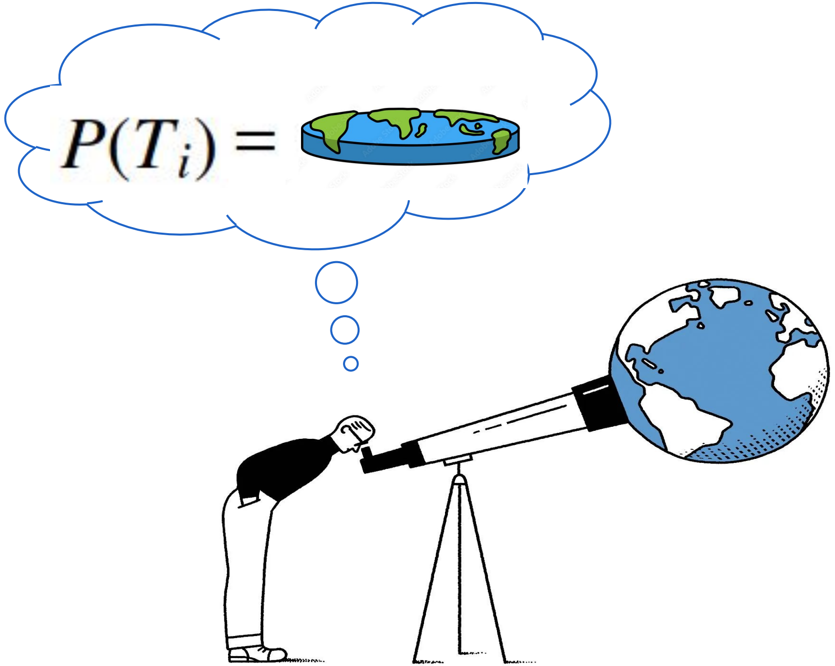
(beauty of bayesian statistics)

→ posterior possibility of various priors converges with sufficient data!

intermezzo 2: humans vs Bayesian principles

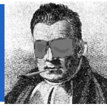


Ideal scenario

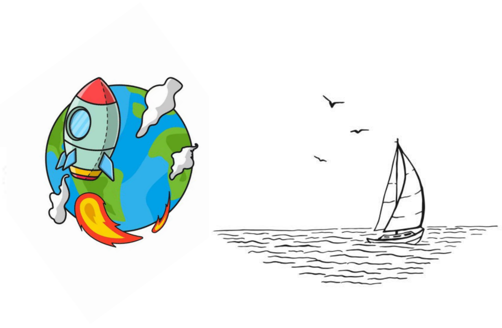


GHENT
UNIVERSITY

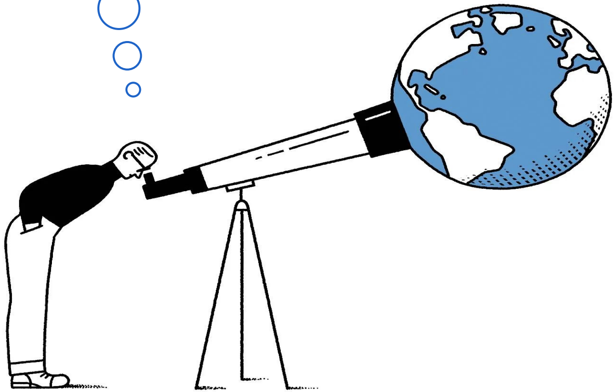
intermezzo 2: humans vs Bayesian principles

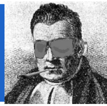


Ideal scenario

$D =$ 

$P(T_i) =$ 

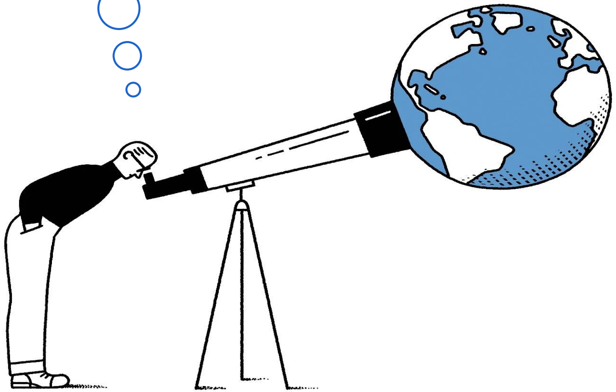


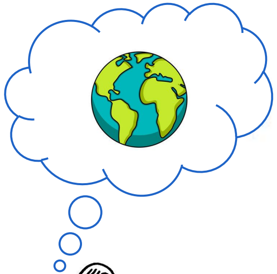


Ideal scenario

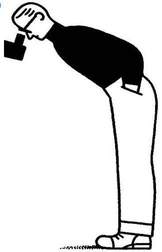
$D =$  

$P(T_i) =$ 

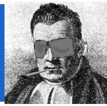


 $= \frac{P(D|T_i)\,P(T_i)}{P(D)}.$

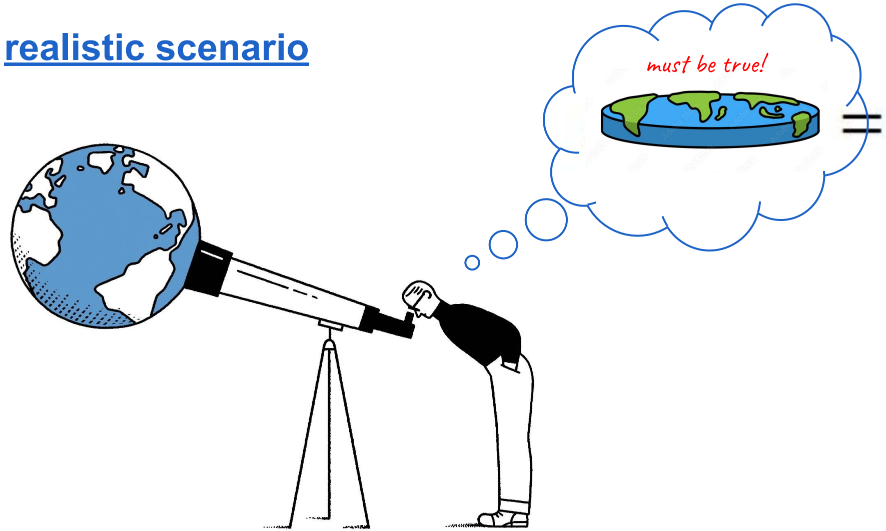
with enough data each prior reaches the same result!



GHENT UNIVERSITY

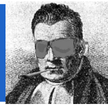


realistic scenario

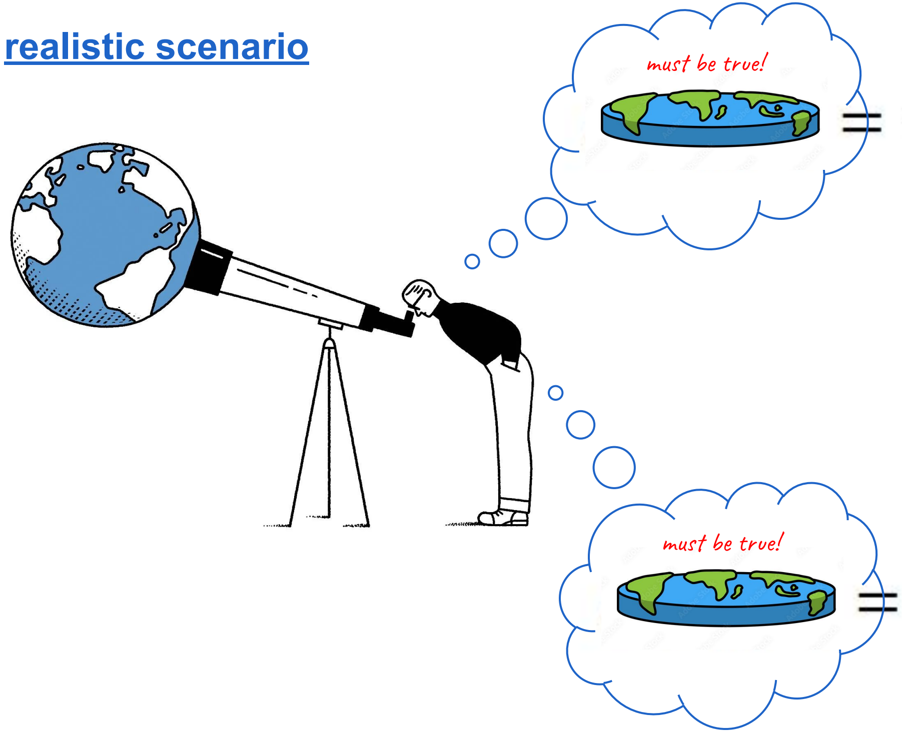


$$= \frac{P(D|T_i)}{P(D)}.$$

intermezzo 2: humans vs Bayesian principles



realistic scenario



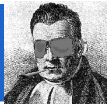
$$= \frac{P(D|T_i)}{P(D)}.$$

new data!

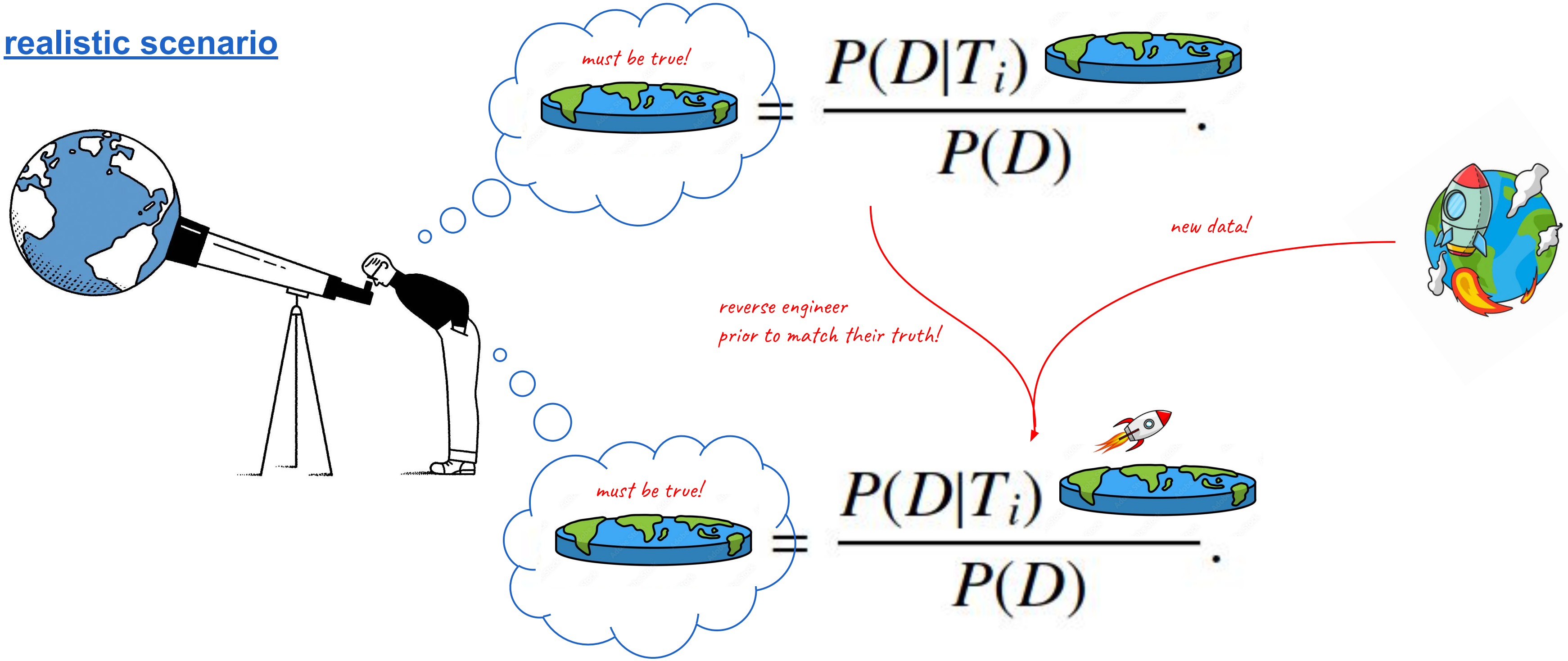
$$= \frac{P(D|T_i)}{P(D)}.$$

GHENT UNIVERSITY

intermezzo 2: humans vs Bayesian principles

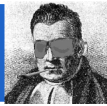


realistic scenario

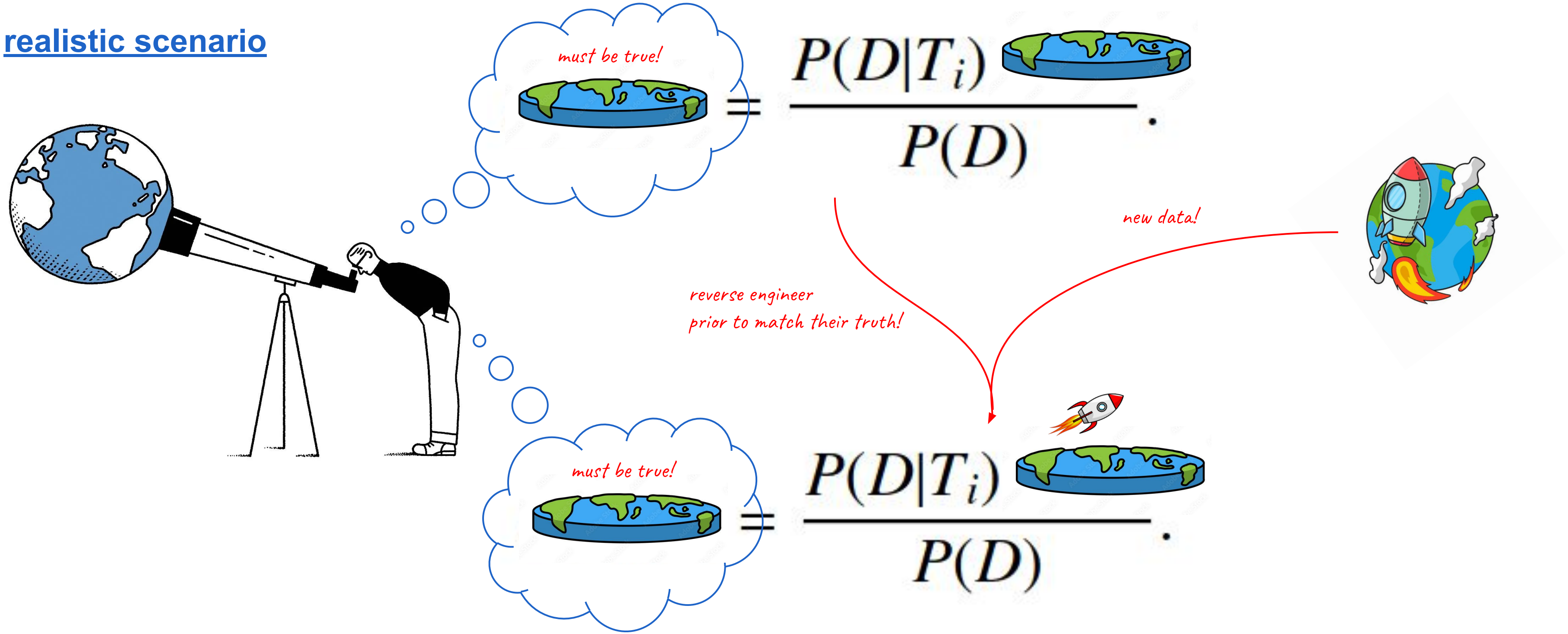


GHENT UNIVERSITY

intermezzo 2: humans vs Bayesian principles



realistic scenario



"In my opinion, we would all reach more reliable conclusions – both individually and as a scientific community – if we were more thoughtful and deliberate in assessing our own Bayesian priors." ~ Dan Hooper

GHENT UNIVERSITY

The Application of the Galactic Center Gamma-Ray Excess

Bayes' theorem applied to his own field!

"I think it is important for experts to clearly state their priors on a variety of propositions – so that others can use that information to guide their own Bayesian reasoning." ~ Dan Hooper

"WHILE DOING THE RESEARCH, KEEP IN MIND THERE ARE ONLY TWO KINDS OF FACTS... THOSE THAT SUPPORT MY POSITION... AND INCONCLUSIVE."

The Application of the Galactic Center Gamma-Ray Excess

Bayes' theorem applied to his own field!

"I think it is important for experts to clearly state their priors on a variety of propositions – so that others can use that information to guide their own Bayesian reasoning." ~ Dan Hooper

SciPost Physics

Submission

The Status of the Galactic Center Gamma-Ray Excess

Dan Hooper

Fermi National Accelerator Laboratory
University of Chicago
* dhooper@fnal.gov

"WHILE DOING THE RESEARCH, KEEP IN MIND THERE ARE ONLY TWO KINDS OF FACTS... THOSE THAT SUPPORT MY POSITION... AND INCONCLUSIVE."

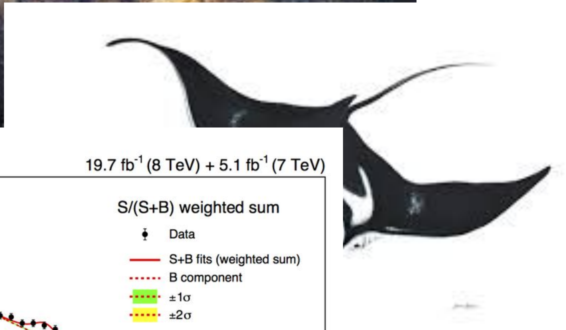
GHENT UNIVERSITY

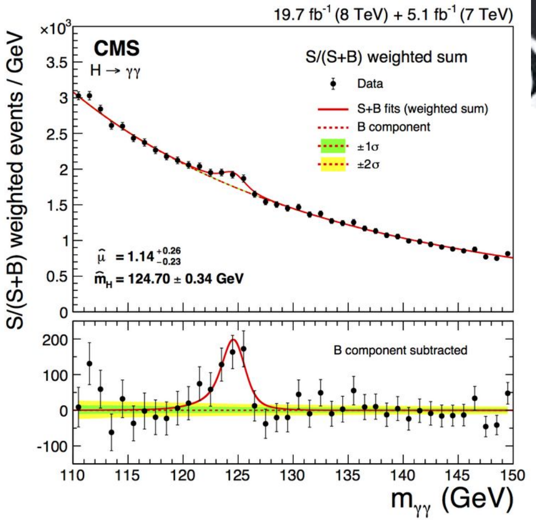
intermezzo 3: The Galactic Center Gamma-Ray Excess

intermezzo 3: The Galactic Center Gamma-Ray Excess

γ-

intermezzo 3: The Galactic Center Gamma-Ray Excess





GHENT UNIVERSITY

intermezzo 3: The Galactic Center Gamma-Ray Excess

- Excess of GeV-scale gamma-ray from the region surrounding the Galactic Center (by Fermi Telescope)

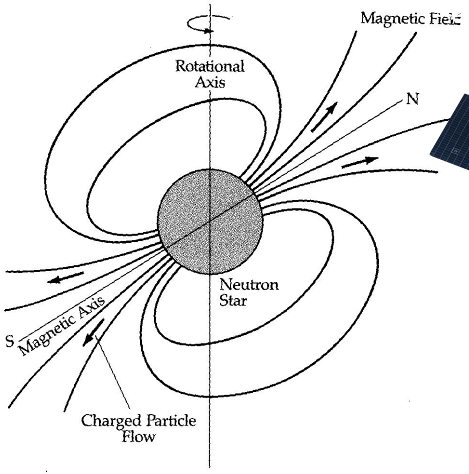
GHENT UNIVERSITY

intermezzo 3: The Galactic Center Gamma-Ray Excess

- Excess of GeV-scale gamma-ray from the region surrounding the Galactic Center (by Fermi Telescope)

<u>large population of millisecond pulsars</u>

→ rotating neutron star with 10 ms rotation period
→ some emit x-ray spectrum



Generic model for a pulsar.

GHENT UNIVERSITY

intermezzo 3: The Galactic Center Gamma-Ray Excess

- Excess of GeV-scale gamma-ray from the region surrounding the Galactic Center (by Fermi Telescope)

large population of millisecond pulsars

→ rotating neutron star with 10 ms rotation period
→ some emit x-ray spectrum

Dark matter annihilation

→ thermal relic (freeze-out)
→m_χ ~ 50 GeV
→$\langle\sigma v\rangle$ ~ 10^{-26} cm^3/s

GHENT UNIVERSITY

The Application of the Galactic Center Gamma-Ray Excess

1) stating priors of the theories considered

(the priors given are largely subjective and for non-astrophysicists feel arbitrary)

2) consider new data to evaluate *P(Data | Theory)*

(probability of the data being measured if a theory is true, not subjective to the scientist that does the measurement)

3) calculate posterior probabilities of the theories

"WHILE DOING THE RESEARCH, KEEP IN MIND THERE ARE ONLY TWO KINDS OF FACTS... THOSE THAT SUPPORT MY POSITION... AND INCONCLUSIVE."

GHENT UNIVERSITY

The Application of the Galactic Center Gamma-Ray Excess

1) stating priors of the theories considered

(the priors given are largely subjective and for non-astrophysicists feel arbitrary)

large population of millisecond pulsars → prior: 10%

- 20% probability that there are enough pulsars to produce a γ-ray signal comparable to the excess
- 50% probability that the pulsars' spatial distribution matches the overserved γ-ray excess

→ Dan's estimate of prior probability!
(no data involved yet)

"WHILE DOING THE RESEARCH, KEEP IN MIND THERE ARE ONLY TWO KINDS OF FACTS... THOSE THAT SUPPORT MY POSITION... AND INCONCLUSIVE."

GHENT UNIVERSITY

The Application of the Galactic Center Gamma-Ray Excess

1) stating priors of the theories considered

(the priors given are largely subjective and for non-astrophysicists feel arbitrary)

large population of millisecond pulsars ⟶ prior: 10%

- 20% probability that there are enough pulsars to produce a γ-ray signal comparable to the excess
- 50% probability that the pulsars' spatial distribution matches the overserved γ-ray excess

Dark matter annihilation signal ⟶ prior: 5.6%

- 40% probability that dark matter is a thermal relic
- 20% probability that the mass of dark matter is ~ 50 GeV
- 70% probability that the annihilation cross section is within a factor 5 range of ~ 10^{-26} cm^3/s

⟶ Dan's estimate of prior probability!
(no data involved yet)

"WHILE DOING THE RESEARCH, KEEP IN MIND THERE ARE ONLY TWO KINDS OF FACTS... THOSE THAT SUPPORT MY POSITION... AND INCONCLUSIVE."

GHENT UNIVERSITY

The Application of the Galactic Center Gamma-Ray Excess

1) stating priors of the theories considered

(the priors given are largely subjective and for non-astrophysicists feel arbitrary)

<u>large population of millisecond pulsars</u> → prior: 10%

- 20% probability that there are enough pulsars to produce a γ-ray signal comparable to the excess
- 50% probability that the pulsars' spatial distribution matches the overserved γ-ray excess

<u>Dark matter annihilation signal</u> → prior: 5.6%

- 40% probability that dark matter is a thermal relic
- 20% probability that the mass of dark matter is ~ 50 GeV
- 70% probability that the annihilation cross section is within a factor 5 range of ~ 10^{-26} cm^3/s

<u>No such signal appears in data</u> → prior: 85%

→ Dan's estimate of prior probability!

(no data involved yet, only assumptions based on previous evidence)

"WHILE DOING THE RESEARCH, KEEP IN MIND THERE ARE ONLY TWO KINDS OF FACTS... THOSE THAT SUPPORT MY POSITION... AND INCONCLUSIVE."

The Application of the Galactic Center Gamma-Ray Excess

2) consider data to evaluate *P(Data | Theory)*

(probability of this data being measured if a theory is true, not subjective to the scientist that does the measurement)

Possible Evidence For Dark Matter Annihilation…

“The analysis performed here would not differentiate the resulting background from dark matter annihilation products. Gamma rays from pion decay taking place with a roughly spherically symmetric distribution around the Galactic Center, for example, could be difficult to distinguish”

$$P(D|T_{\mathrm{DM}}) \sim P(D|T_{\mathrm{pulsars}}) \sim 1 \qquad P(D|T_{\mathrm{Neither}}) \sim 0.$$

"WHILE DOING THE RESEARCH, KEEP IN MIND THERE ARE ONLY TWO KINDS OF FACTS... THOSE THAT SUPPORT MY POSITION... AND INCONCLUSIVE."

The Application of the Galactic Center Gamma-Ray Excess

2) consider data to evaluate *P(Data|Theory)*

(probability of this data being measured if a theory is true, not subjective to the scientist that does the measurement)

Possible Evidence For Dark Matter Annihilation…

"The analysis performed here would not differentiate the resulting background from dark matter annihilation products. Gamma rays from pion decay taking place with a roughly spherically symmetric distribution around the Galactic Center, for example, could be difficult to distinguish"

$$P(D|T_{\mathrm{DM}}) \sim P(D|T_{\mathrm{pulsars}}) \sim 1 \qquad P(D|T_{\mathrm{Neither}}) \sim 0.$$

3) calculate posterior probabilities of the theories

$$P(T_{\mathrm{DM}}|D) \sim 0.36,\ P(T_{\mathrm{pulsars}}|D) \sim 0.64, \text{ and } P(T_{\mathrm{neither}}|D) \sim 0.$$

"WHILE DOING THE RESEARCH, KEEP IN MIND THERE ARE ONLY TWO KINDS OF FACTS... THOSE THAT SUPPORT MY POSITION... AND INCONCLUSIVE."

GHENT UNIVERSITY

The Application of the Galactic Center Gamma-Ray Excess

2) consider data to evaluate *P(Data□Theory)*

(probability of this data being measured if a theory is true, not subjective to the scientist that does the measurement)

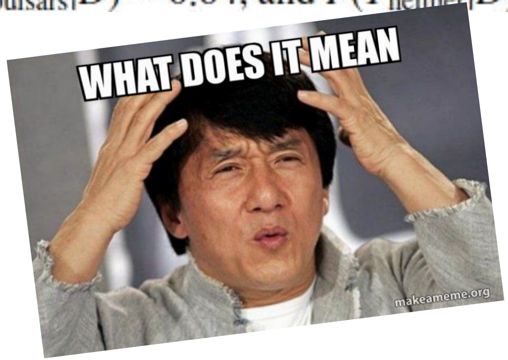
Possible Evidence For Dark Matter Annihilation…

“The analysis performed here would not differentiate the resulting background from dark matter annihilation products. Gamma rays from pion decay taking place with a roughly spherically symmetric distribution around the Galactic Center, for example, could be difficult to distinguish”

$$P(D|T_{\mathrm{DM}}) \sim P(D|T_{\mathrm{pulsars}}) \sim 1 \qquad P(D|T_{\mathrm{Neither}}) \sim 0.$$

3) calculate posterior probabilities of the theories

$$P(T_{\mathrm{DM}}|D) \sim 0.36,\ P(T_{\mathrm{pulsars}}|D) \sim 0.64, \text{ and } P(T_{\mathrm{neither}}|D) \sim 0.$$



“WHILE DOING THE RESEARCH, KEEP IN MIND THERE ARE ONLY TWO KINDS OF FACTS… THOSE THAT SUPPORT MY POSITION… AND INCONCLUSIVE.”

GHENT UNIVERSITY

What about a discovery in high energy physics?

"In the field of particle physics, the word discovery is sometimes treated as a synonym for one or more measurements that are discrepant with the predictions of the null hypothesis at a level of at least 5 standard deviations."

- "look elsewhere" effect
- plausibility of the searched new physics
- role of systematics
- …

not all searches for new physics at the LHC are the same in these aspects![4]

→ should each search for new physics be treated with the same precision?
→ can the frequentist and bayesian approach be combined somehow?
→ can we step away from discovery in general? (just mention the p-value)

conclusions?

a discovery does not necessarily require a 5-σ posterior probability!

Then what is the use?

- includes social nature of human assessment of experimental anomalies
- can quantify why previous examples are not considered a discovery

conclusions?

a discovery does not necessarily require a 5-σ posterior probability!

Then what is the use?

- includes social nature of human assessment of experimental anomalies
- can quantify why previous examples are not considered a discovery

What goes into the response of the scientific community to new data and experimental anomalies?

> “We are all more skeptical of information that surprises us, and are more likely to accept information that confirms our expectations. This is Bayesian inference at work.”

GHENT UNIVERSITY

conclusions?

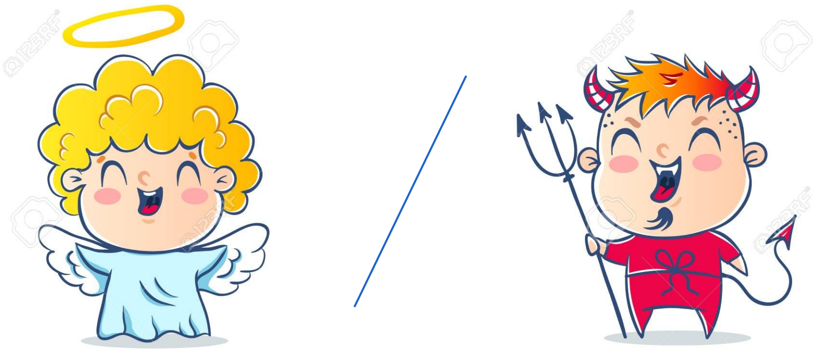
a discovery does not necessarily require a 5-σ posterior probability!

Then what is the use?

- includes social nature of human assessment of experimental anomalies
- can quantify why previous examples are not considered a discovery

What goes into the response of the scientific community to new data and experimental anomalies?

“We are all more skeptical of information that surprises us, and are more likely to accept information that confirms our expectations. This is Bayesian inference at work.”



GHENT UNIVERSITY

FACULTY OF SCIENCES

Discussion!

GHENT UNIVERSITY

references

[1] New results from DAMA/LIBRA, R. Bernabei et. al.
[2] New results from the Muon g-2 Experiments, G. Venanzoni (Muon g-2 Collaboration)
[3] Trial factors for the look elsewhere effect in high energy physics, E. Gross, and O. Vitels
[4] DISCOVERING THE SIGIFICANCE OF 5σ, L. Lyons

Podcast on the Galactic Center Gamma-Ray Excess by Dan Hooper

Music by Dan’s university physics-punk band (him included)

GHENT UNIVERSITY