

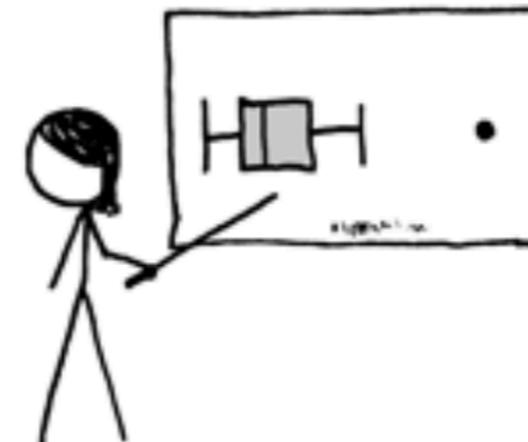
CAN MY BOYFRIEND
COME ALONG?



I'M NOT YOUR
BOYFRIEND!
I YOU TOTALLY ARE.
I'M CASUALLY
DATING A NUMBER
OF PEOPLE.

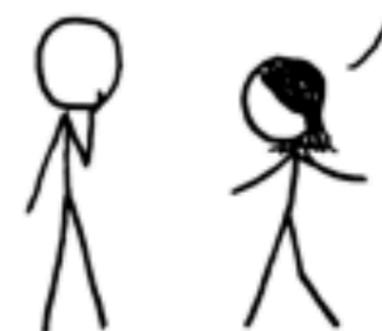


BUT YOU SPEND TWICE AS MUCH
TIME WITH ME AS WITH ANYONE
ELSE. I'M A CLEAR OUTLIER.



YOUR MATH IS
IRREFUTABLE.

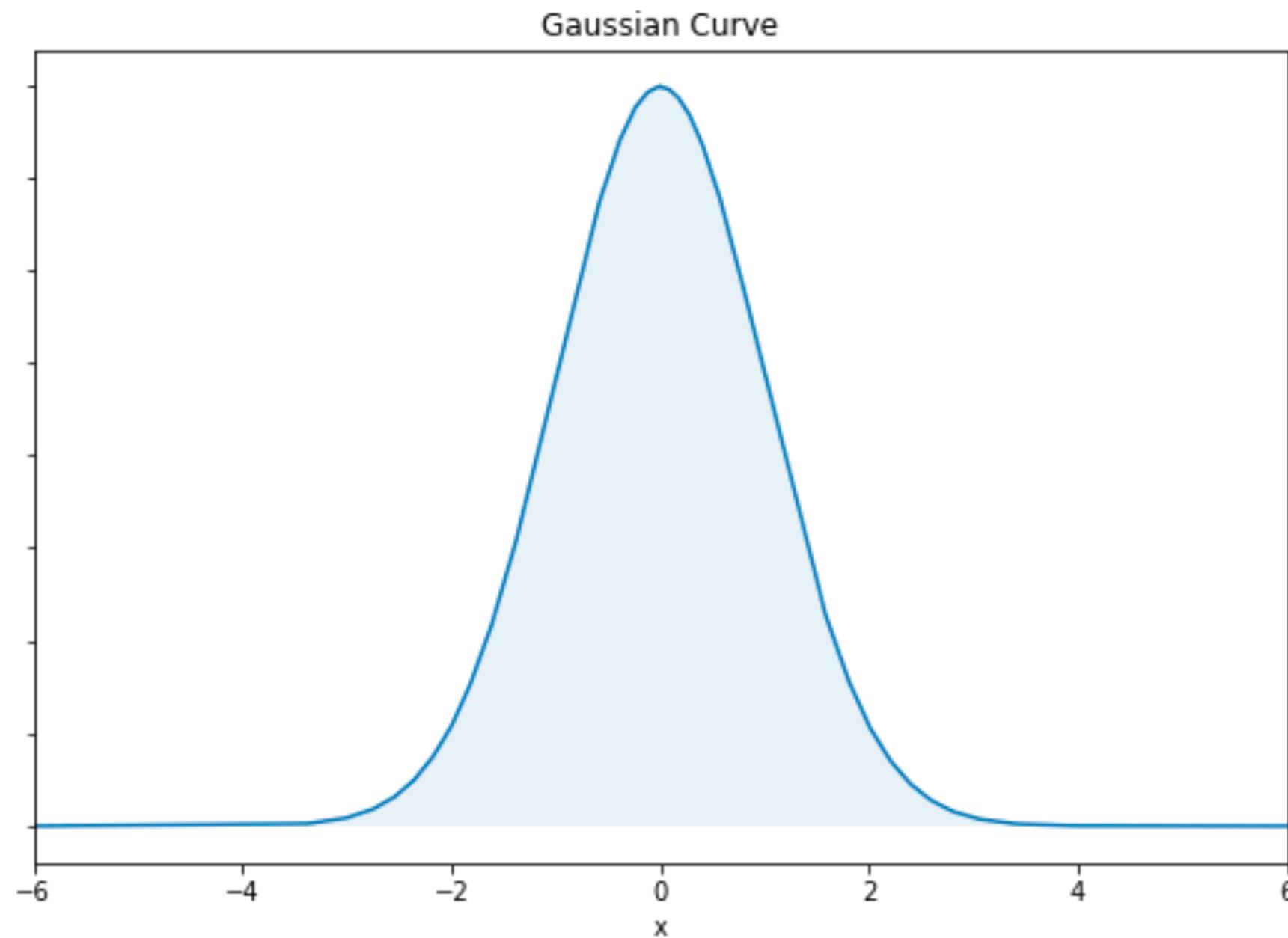
FACE IT—I'M
YOUR STATISTICALLY
SIGNIFICANT OTHER.



Lecture 9: Hypothesis Testing

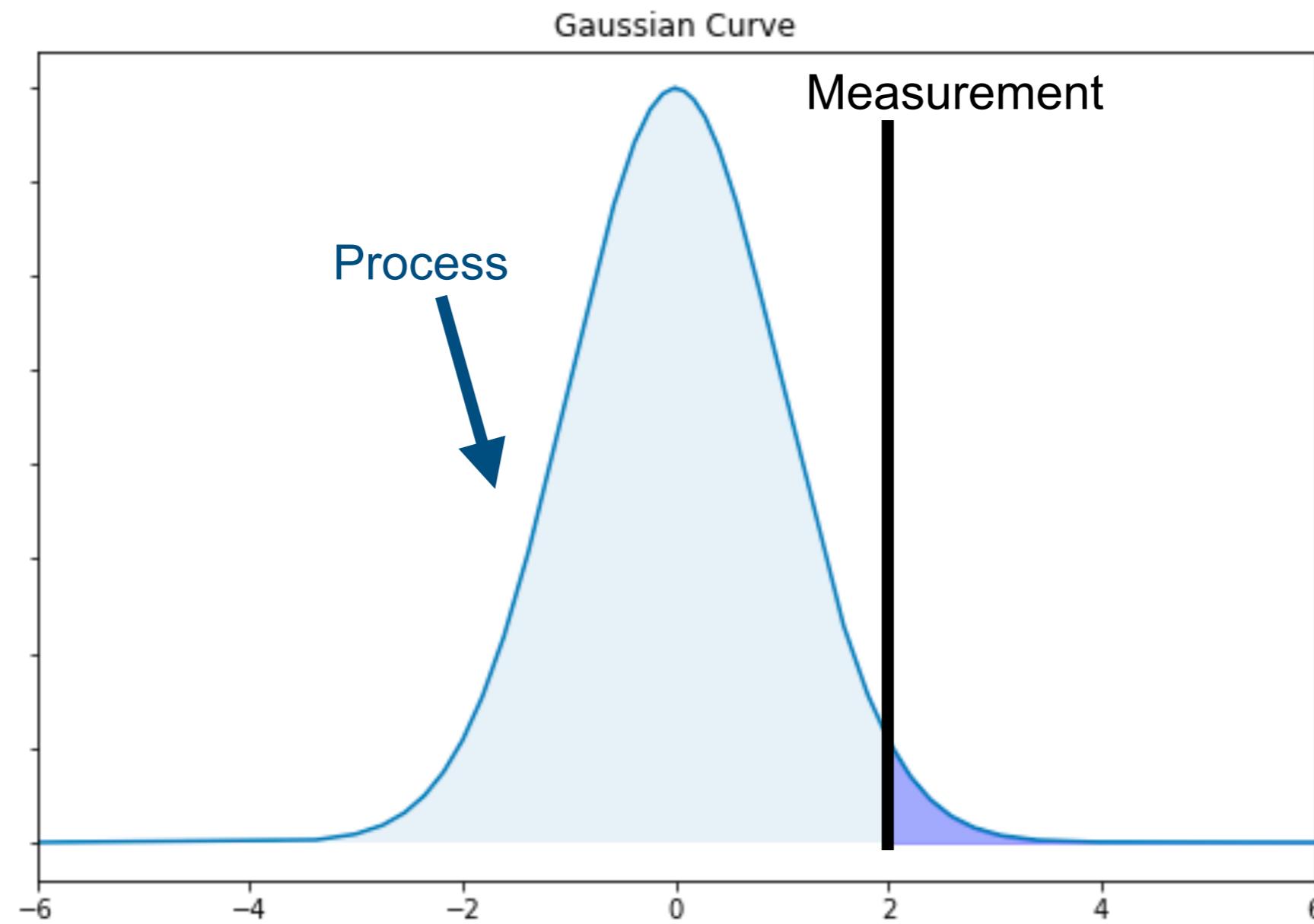
What is a measurement?

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



What is a measurement?

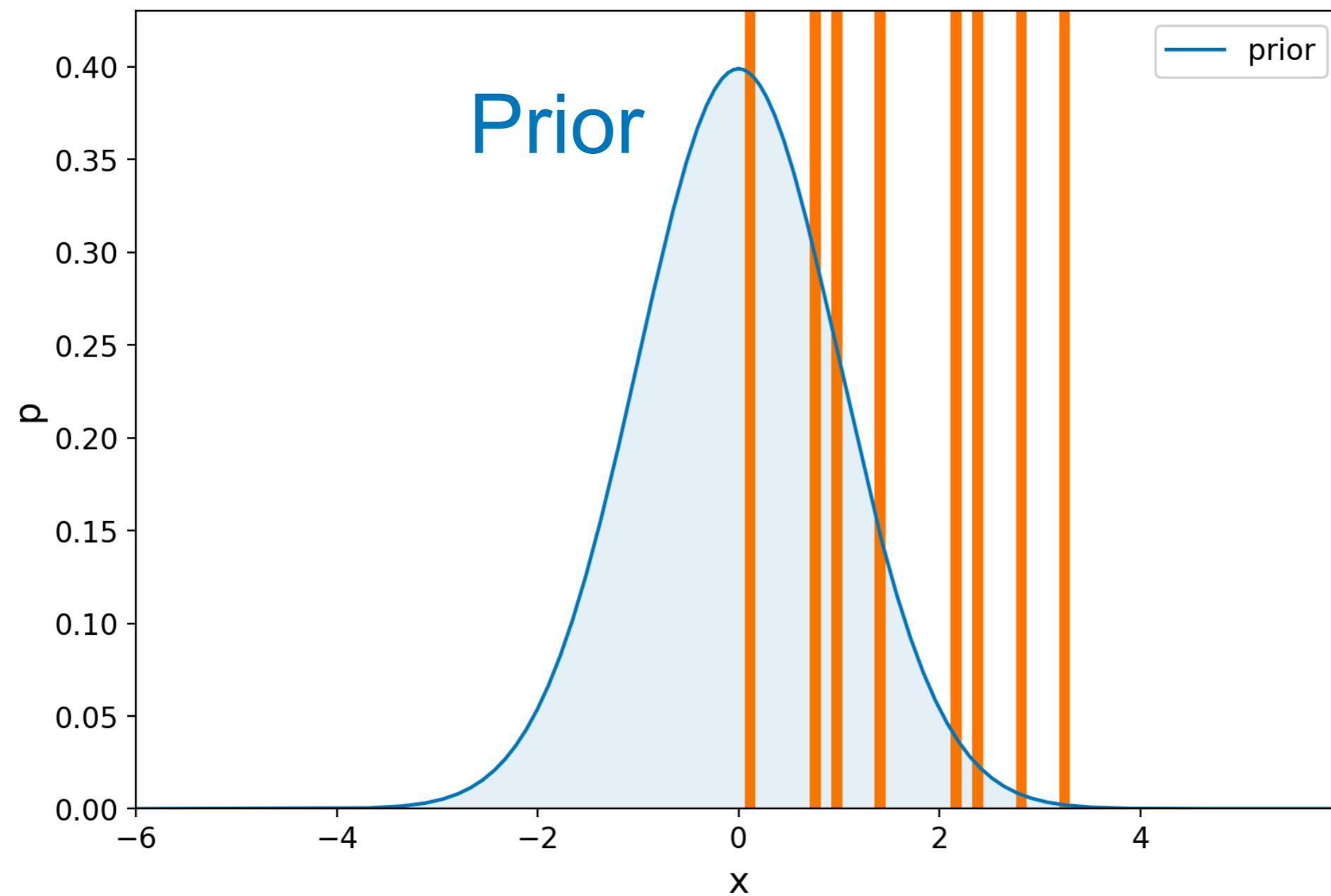
- We have assumed a model for this process
 - Example what is the direction of your eye reading slide?



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

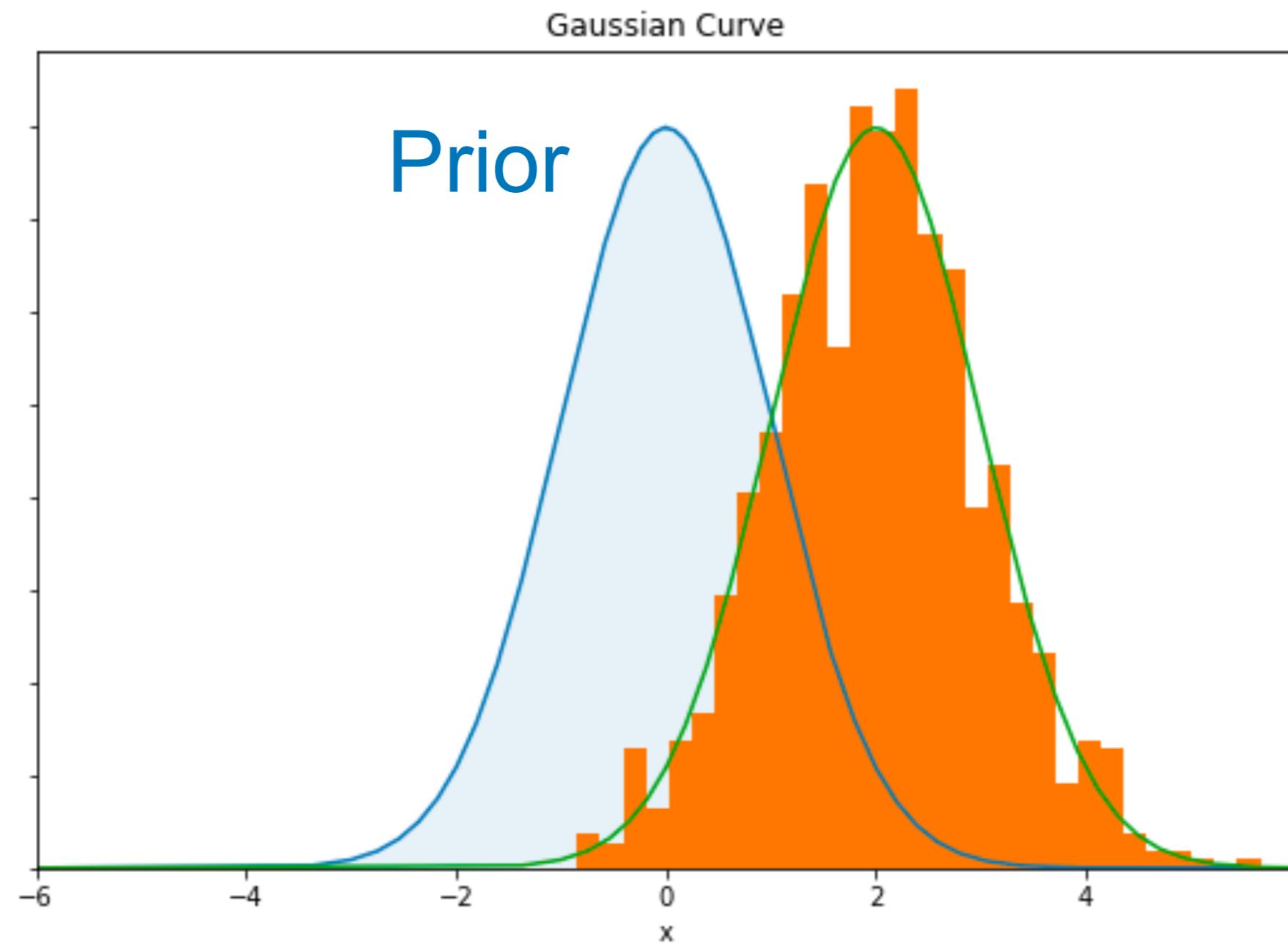
What happens with many⁴ measurements?

- What if the distribution we observe is different?



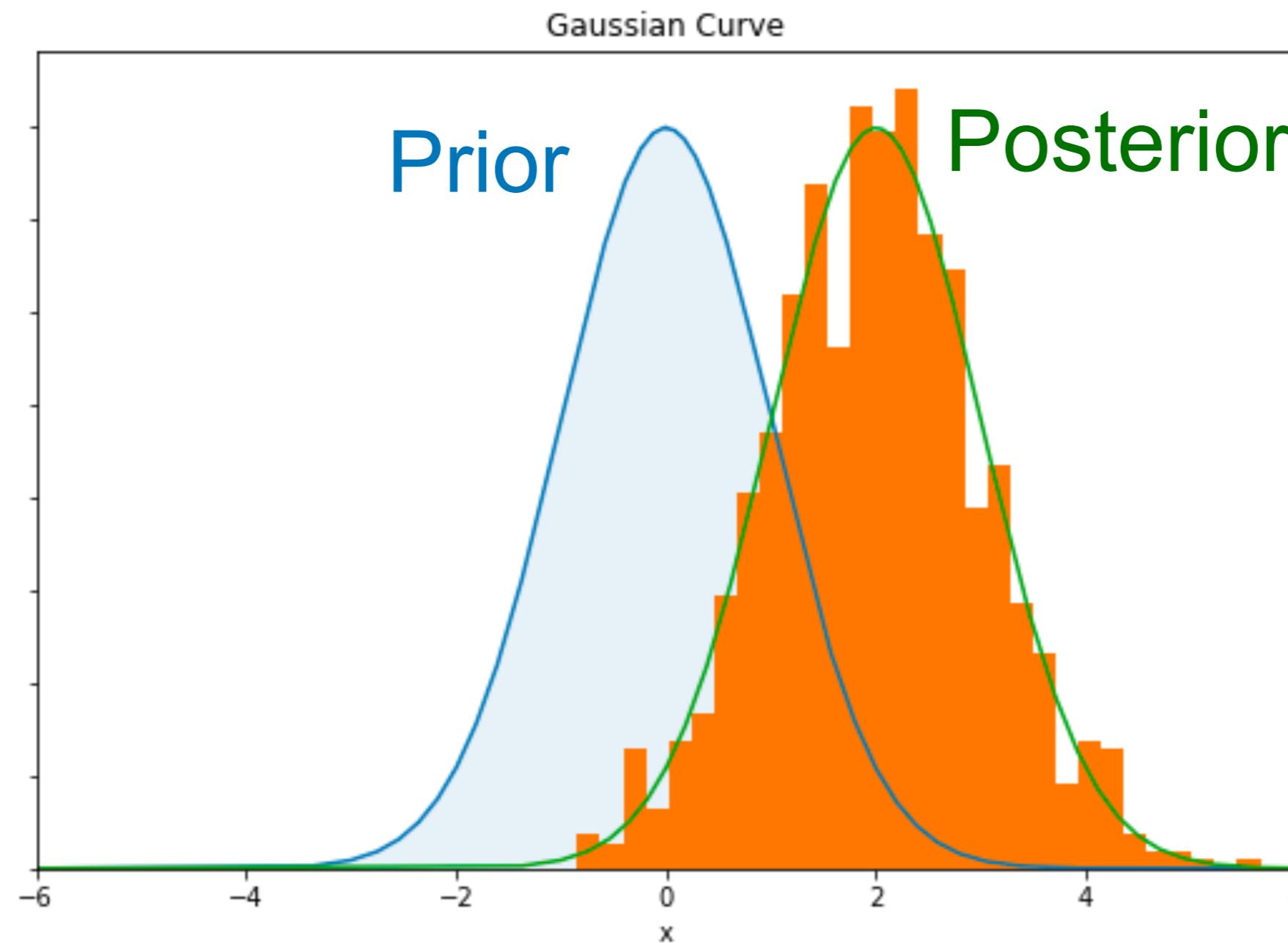
What happens with many⁵ measurements?

- What if the distribution we observe is different?

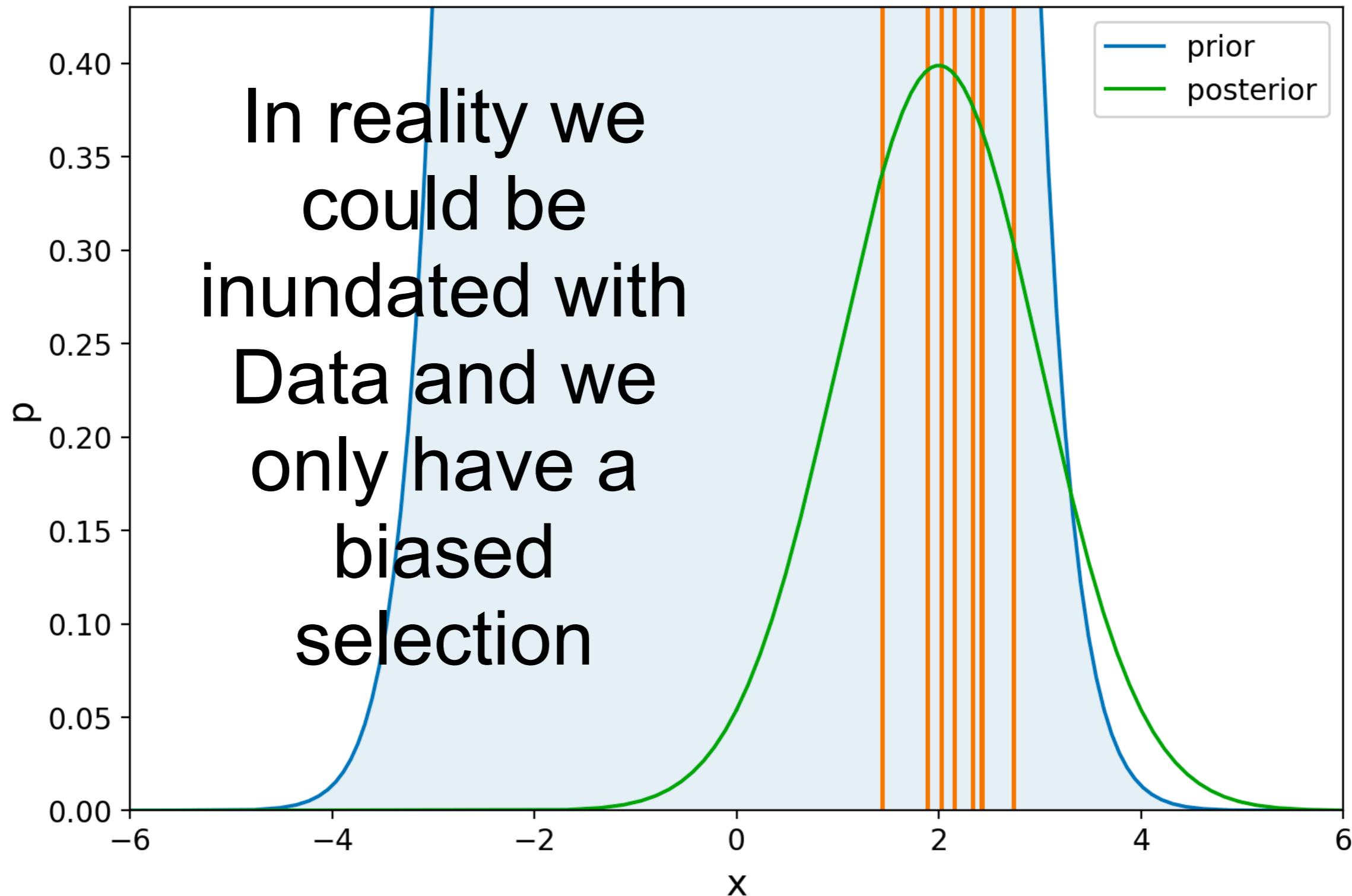


What happens with many⁶ measurements?

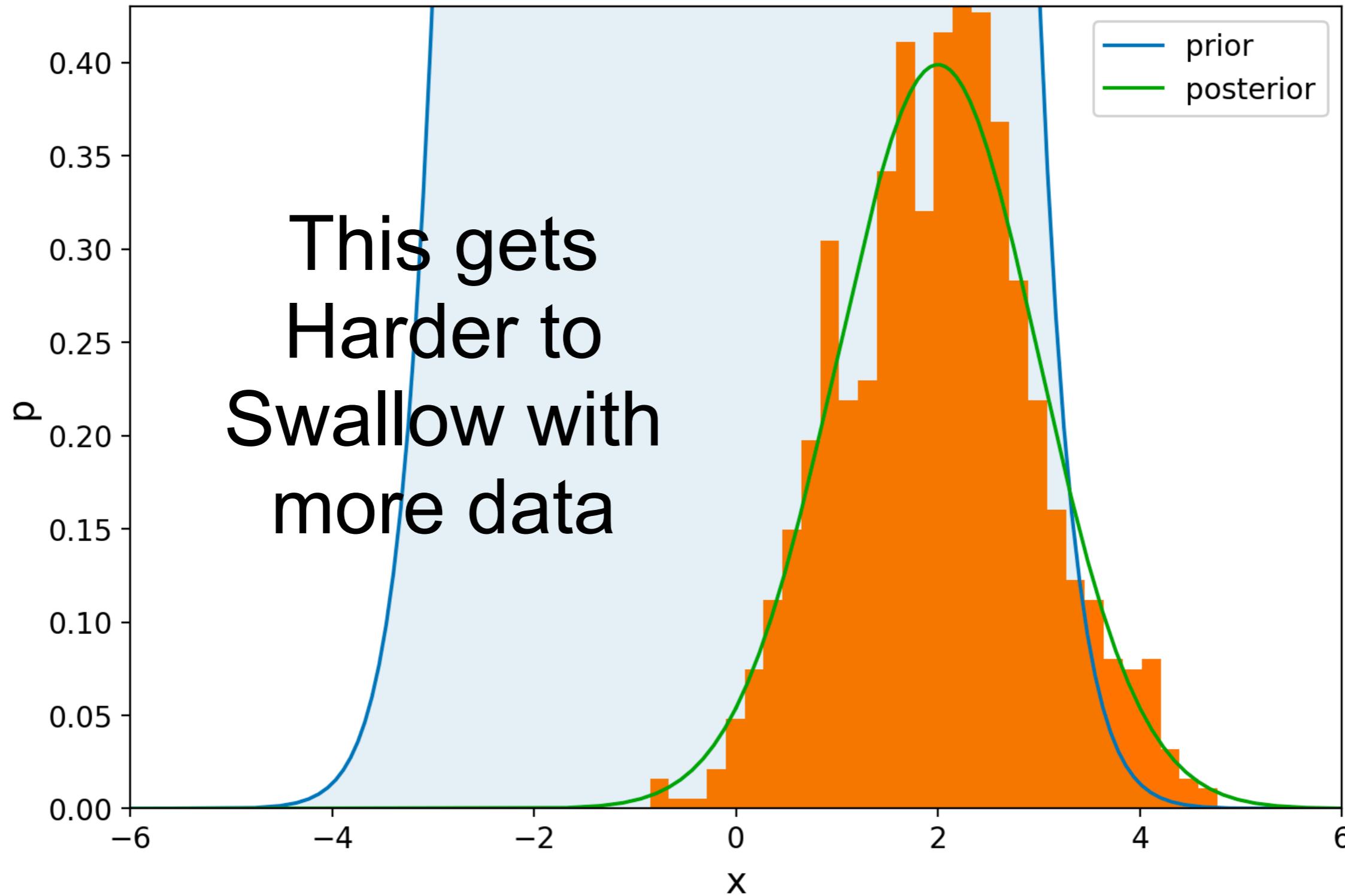
- What if the distribution we observe is different?



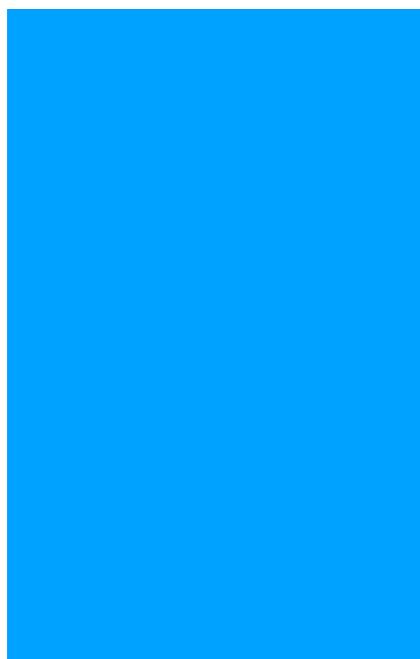
Are we missing something?



Are we missing⁸ something?



Monty Hall Problem



Door 1



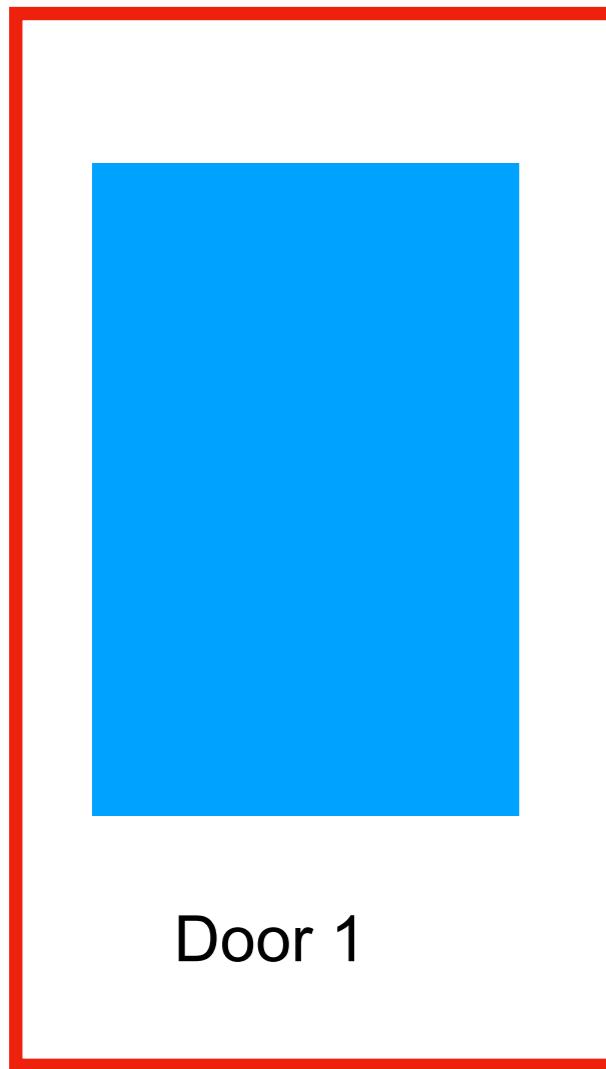
Door 2



Door 3

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

Monty Hall Problem



Door 1



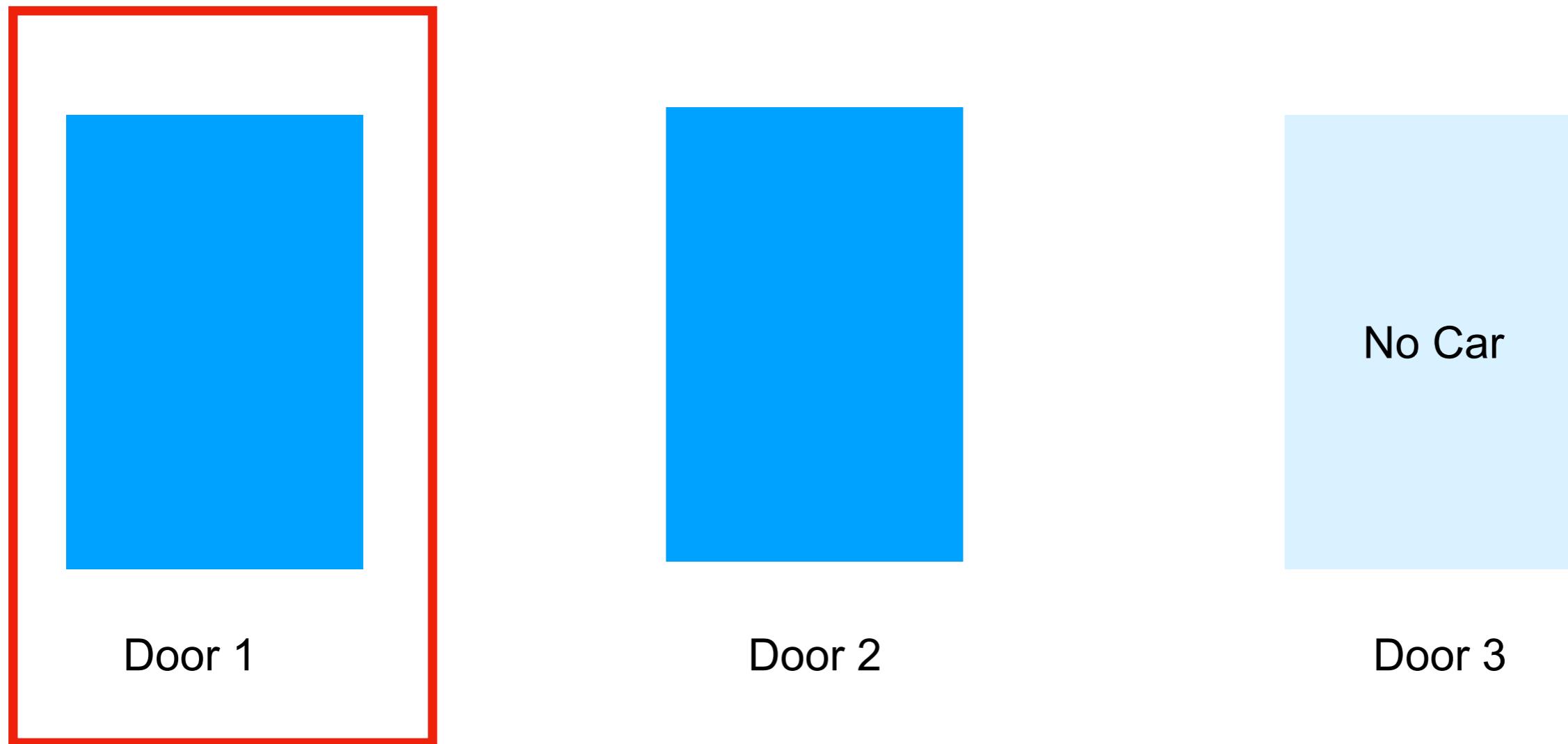
Door 2



Door 3

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 2

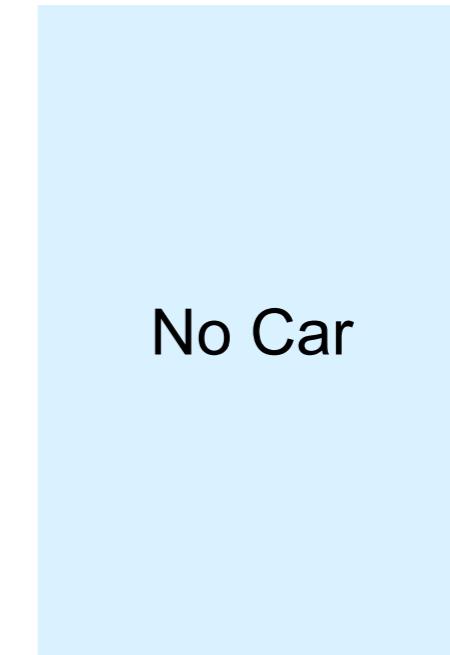
Monty Hall Problem



Door 1



Door 2

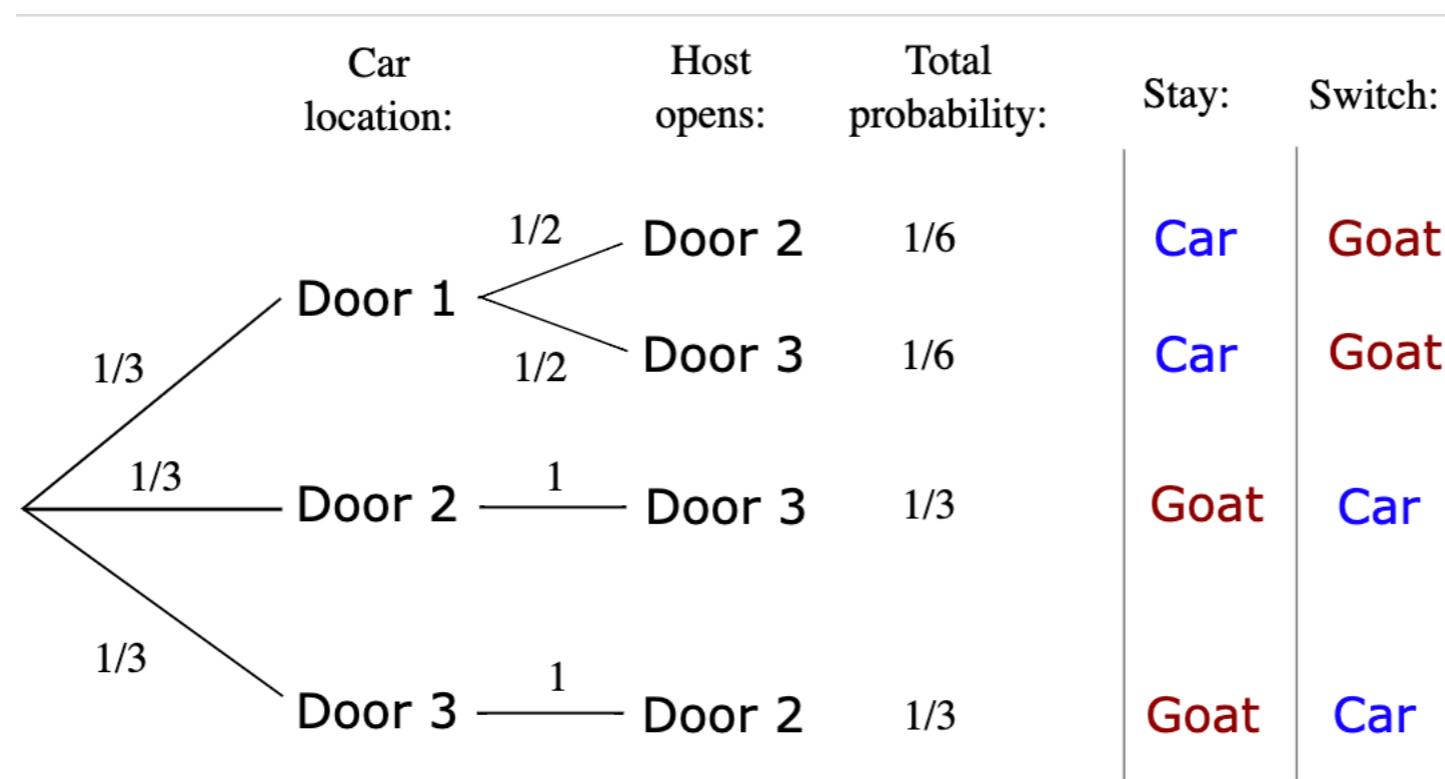


Door 3

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

Prior: Probability of car behind doors P(Car@...)	Event: Probability of Monty to open door B (you chose door A) P(Open B Car@...)*	Posterior Probability : chances of the car behind the doors after the event P(Car@... Opened B)
$P(\text{Car}@A) = \frac{1}{3}$	* Monty will never open a door concealing the car	$P(\text{Car}@A \text{Open } B) = \frac{\left(\frac{1}{2} \times \frac{1}{3}\right)}{\left(\frac{1}{2} \times \frac{1}{3} + 0 \times \frac{1}{3} + 1 \times \frac{1}{3}\right)} = \frac{1}{3}$
$P(\text{Car}@B) = \frac{1}{3}$	$P(\text{Open } B \text{Car}@A) = \frac{1}{2}$	$P(\text{Car}@B \text{Open } B) = \frac{(0 \times \frac{1}{3})}{\left(\frac{1}{2} \times \frac{1}{3} + 0 \times \frac{1}{3} + 1 \times \frac{1}{3}\right)} = 0$
$P(\text{Car}@C) = \frac{1}{3}$	$P(\text{Open } B \text{Car}@B) = 0$	$P(\text{Car}@C \text{Open } B) = \frac{\left(1 \times \frac{1}{3}\right)}{\left(\frac{1}{2} \times \frac{1}{3} + 0 \times \frac{1}{3} + 1 \times \frac{1}{3}\right)} = \frac{2}{3}$

What happens with many¹⁵ measurements?

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

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

Prior

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

Likelihood

Posterior

What happens with many¹⁶ measurements?

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

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

Prior

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

Likelihood

Posterior Normalizer

posterior \propto likelihood \times prior

What happens with many¹⁷ measurements?

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

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

Prior

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

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 | Advanced

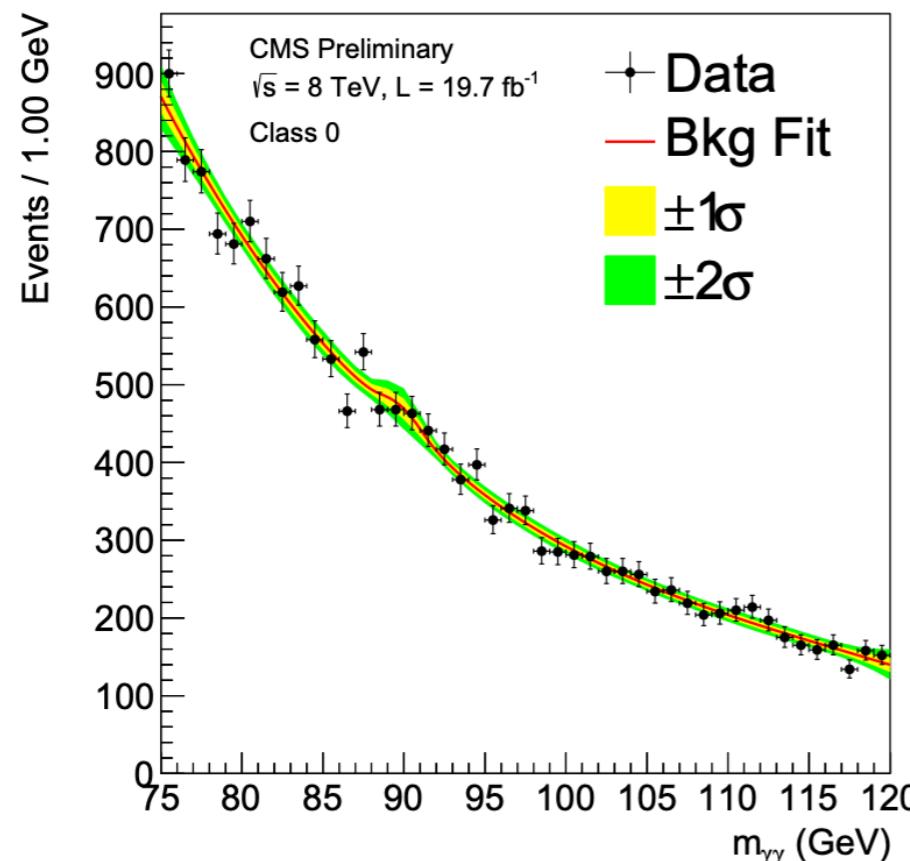
High Energy Physics – Phenomenology

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

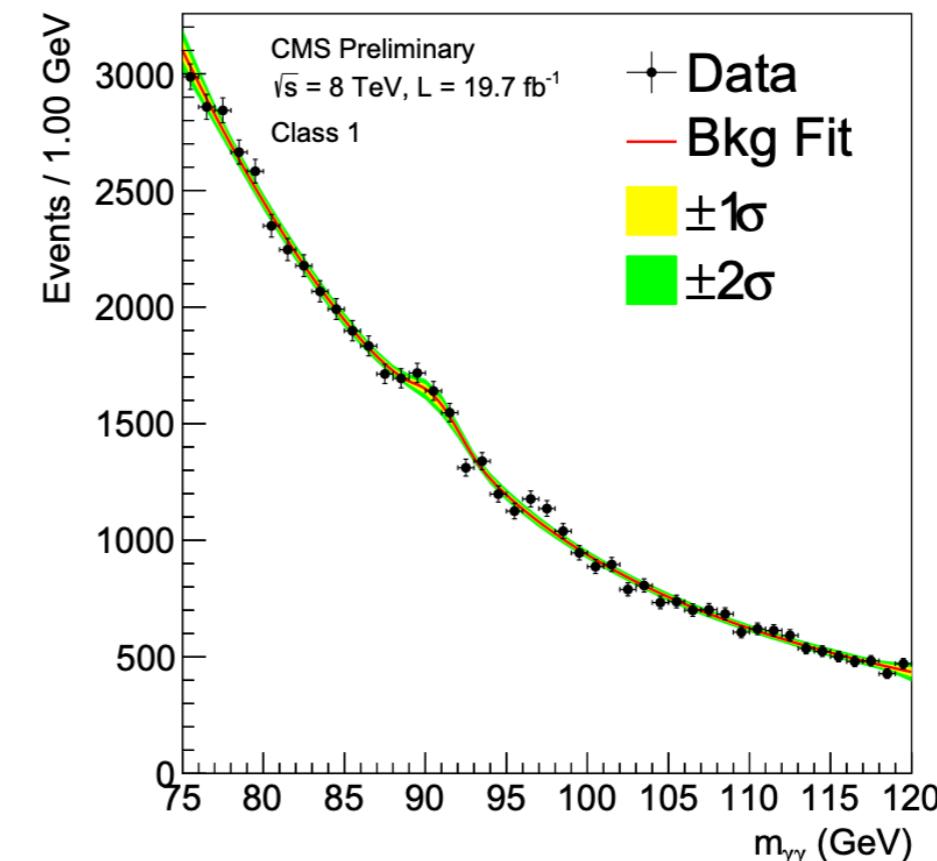
Mounting evidence for a 95 GeV Higgs boson

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 2.8σ . 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 3.1σ 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 2.3σ 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 | Advanced

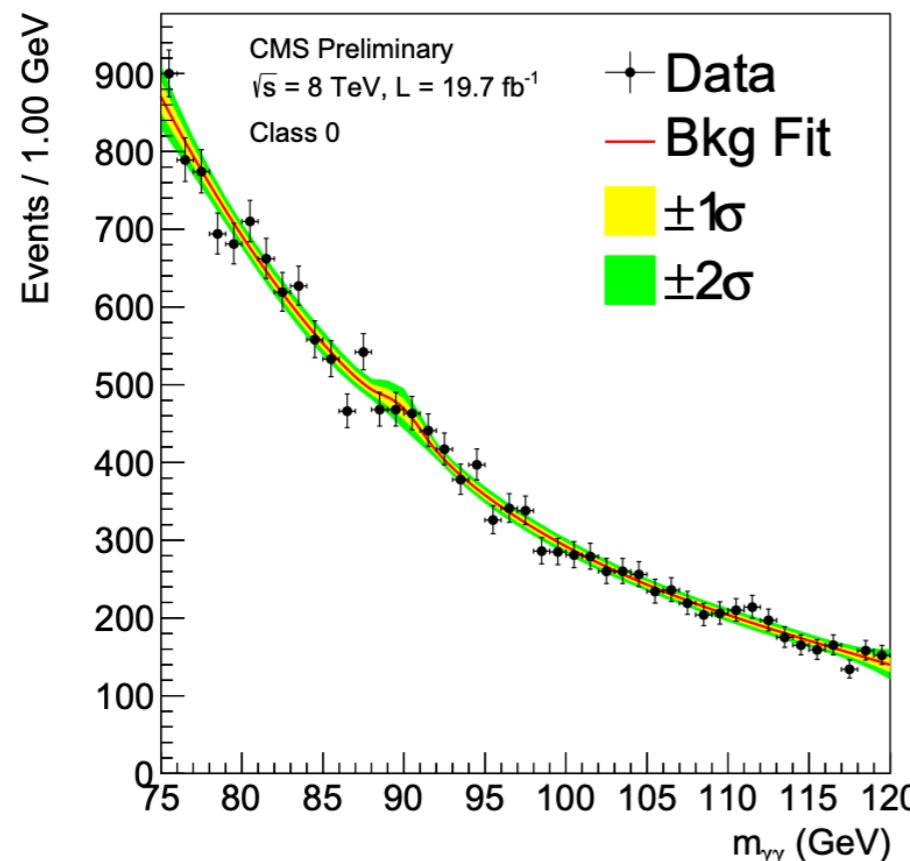
High Energy Physics – Phenomenology

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

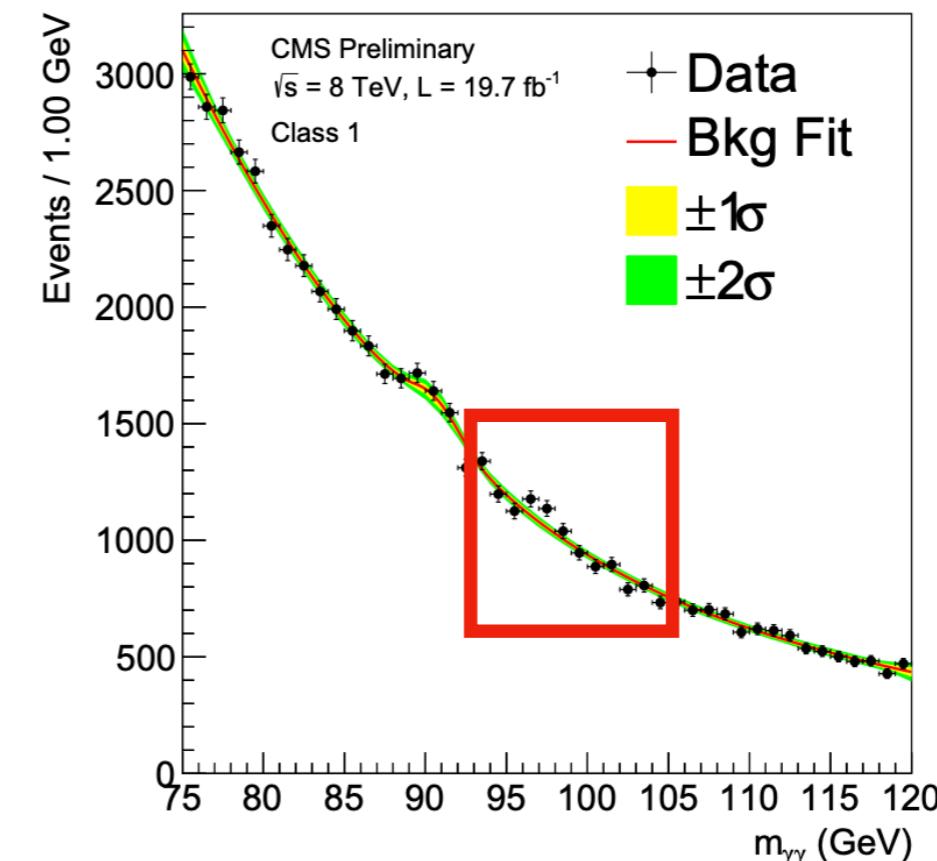
Mounting evidence for a 95 GeV Higgs boson

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 2.8σ . 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 3.1σ 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 2.3σ 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

DID THE SUN JUST EXPLODE?
(IT'S NIGHT, SO WE'RE NOT SURE.)

THIS NEUTRINO DETECTOR MEASURES
WHETHER THE SUN HAS GONE NOVA.

THEN, IT ROLLS TWO DICE. IF THEY
BOTH COME UP SIX, IT LIES TO US.
OTHERWISE, IT TELLS THE TRUTH.

LET'S TRY.

DETECTOR! HAS THE
SUN GONE NOVA?

ROLL

YES.



FREQUENTIST STATISTICIAN:

THE PROBABILITY OF THIS RESULT
HAPPENING BY CHANCE IS $\frac{1}{36} = 0.027$.
SINCE $p < 0.05$, I CONCLUDE
THAT THE SUN HAS EXPLODED.



BAYESIAN STATISTICIAN:

BET YOU \$50
IT HASN'T.



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

**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.

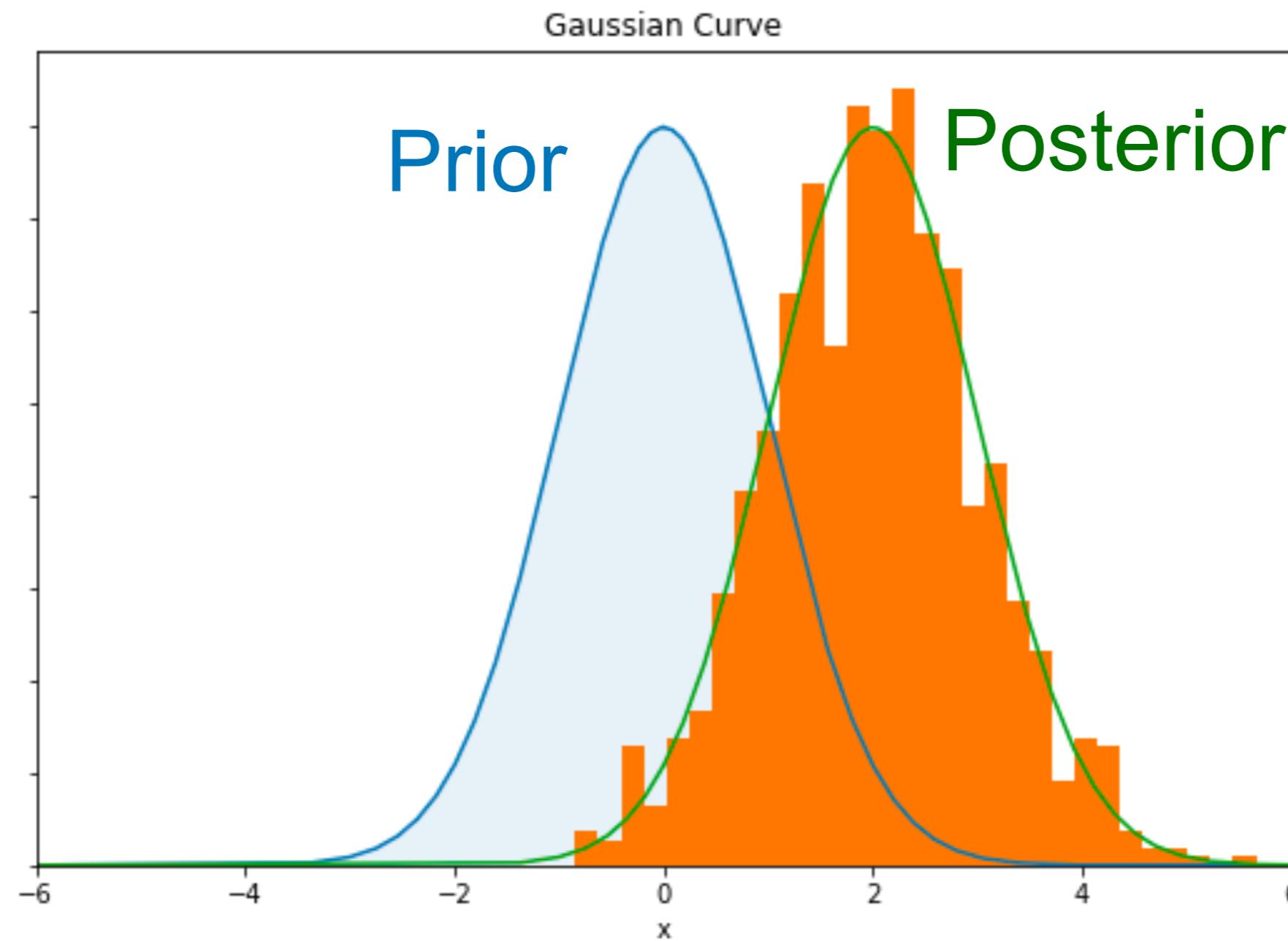
Frequentist path

$$\text{Likelihood } L(H; D) = P(D|H)$$

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

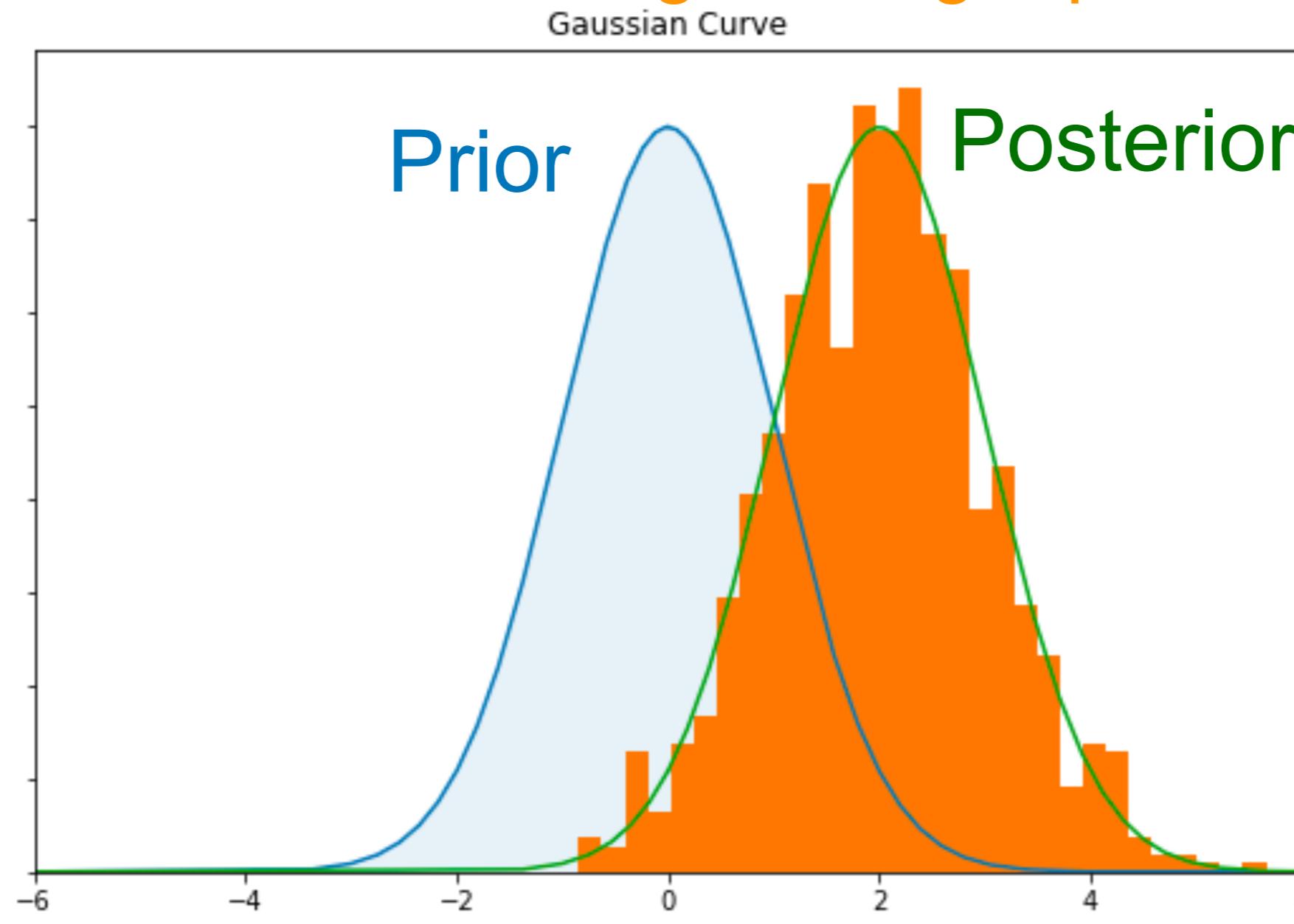
Bayesian View

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



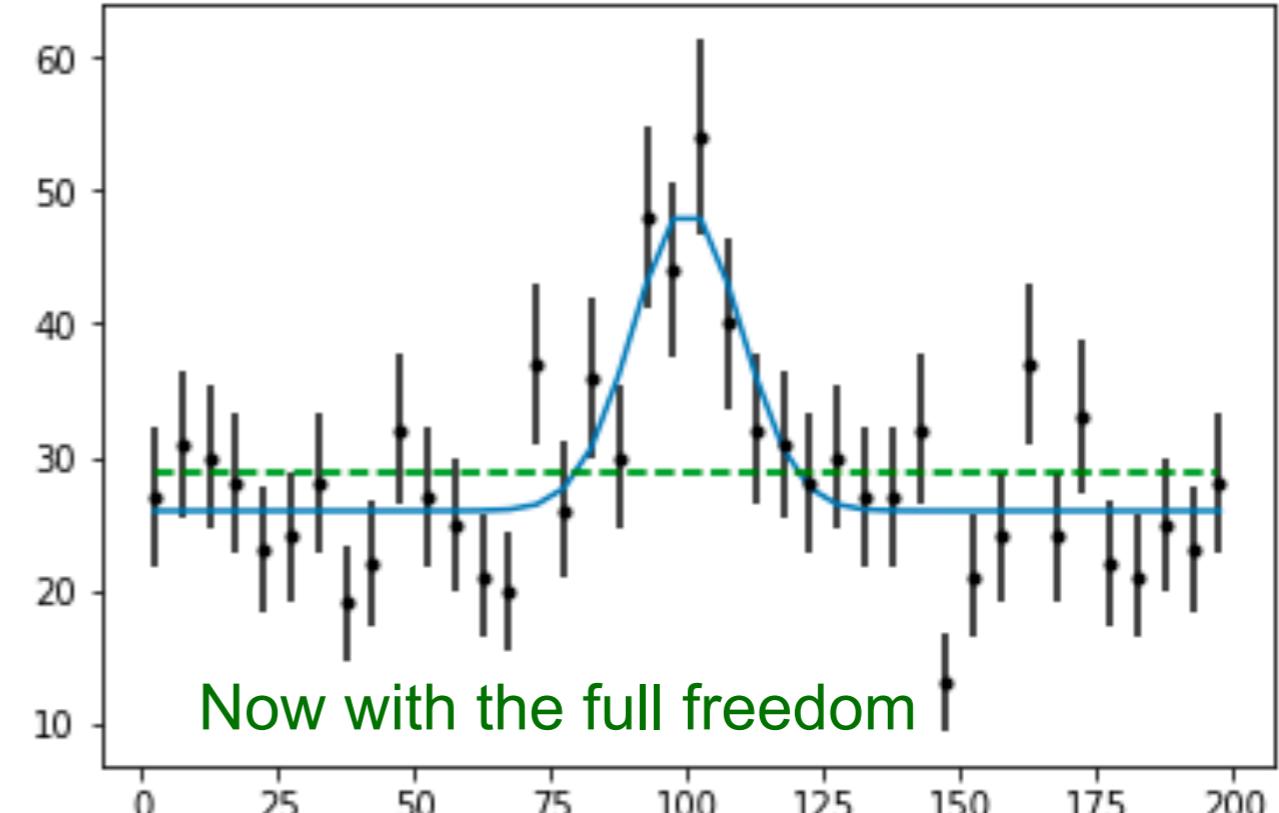
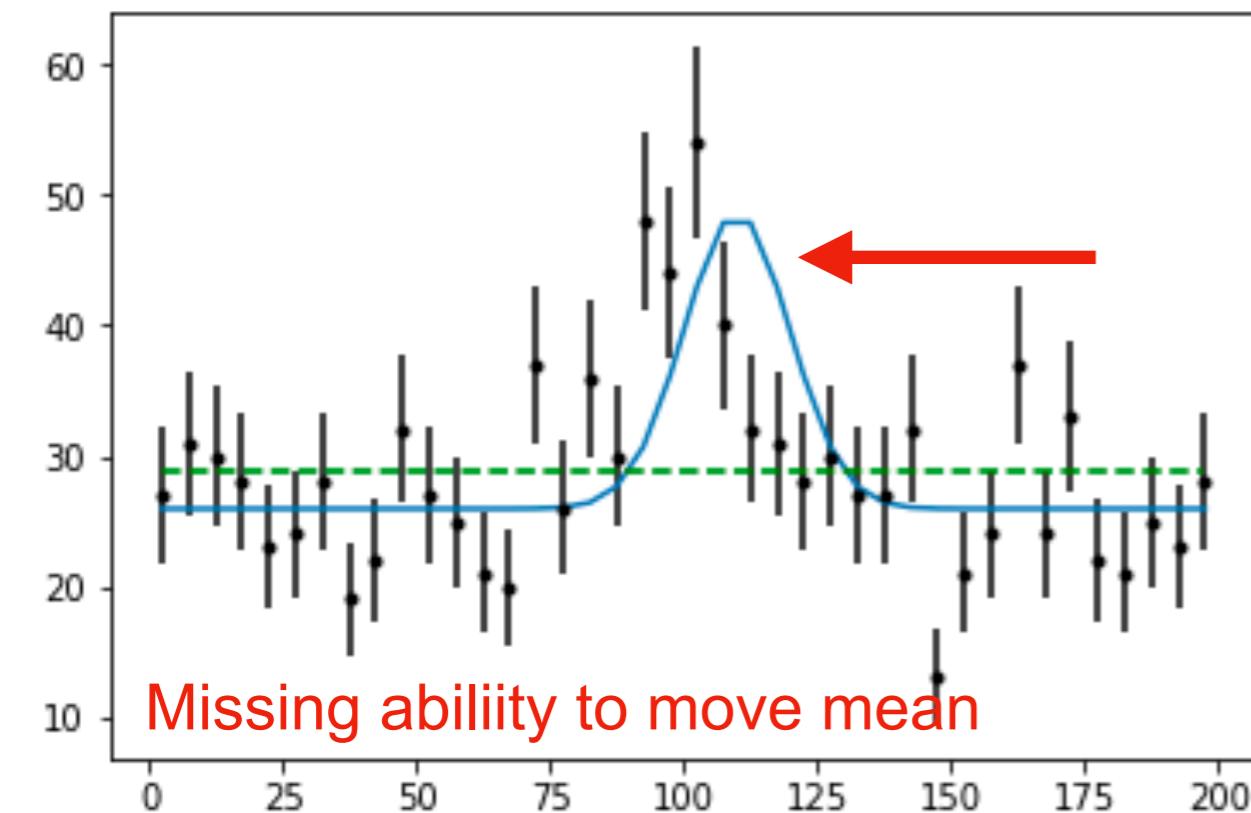
Frequentist View

All we know is our result gives us a gaussian at 2
Bayesian stats people are a bunch of
smart asses guessing a prior



What if models is²⁴ incomplete?

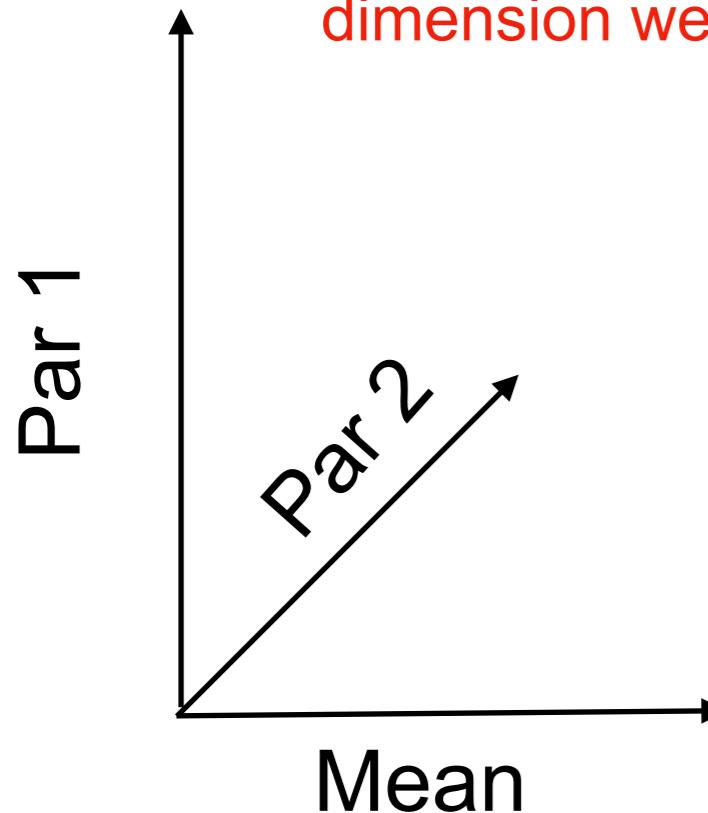
- <https://www.youtube.com/watch?v=Ku42lszh9KM>
- Sometimes there are other parameters not included in our model



What if models is²⁵ incomplete?

- <https://www.youtube.com/watch?v=Ku42lszh9KM>
- Sometimes there are other parameters not included in our model

Bayesian view:
Each parameter adds a
dimension we probe

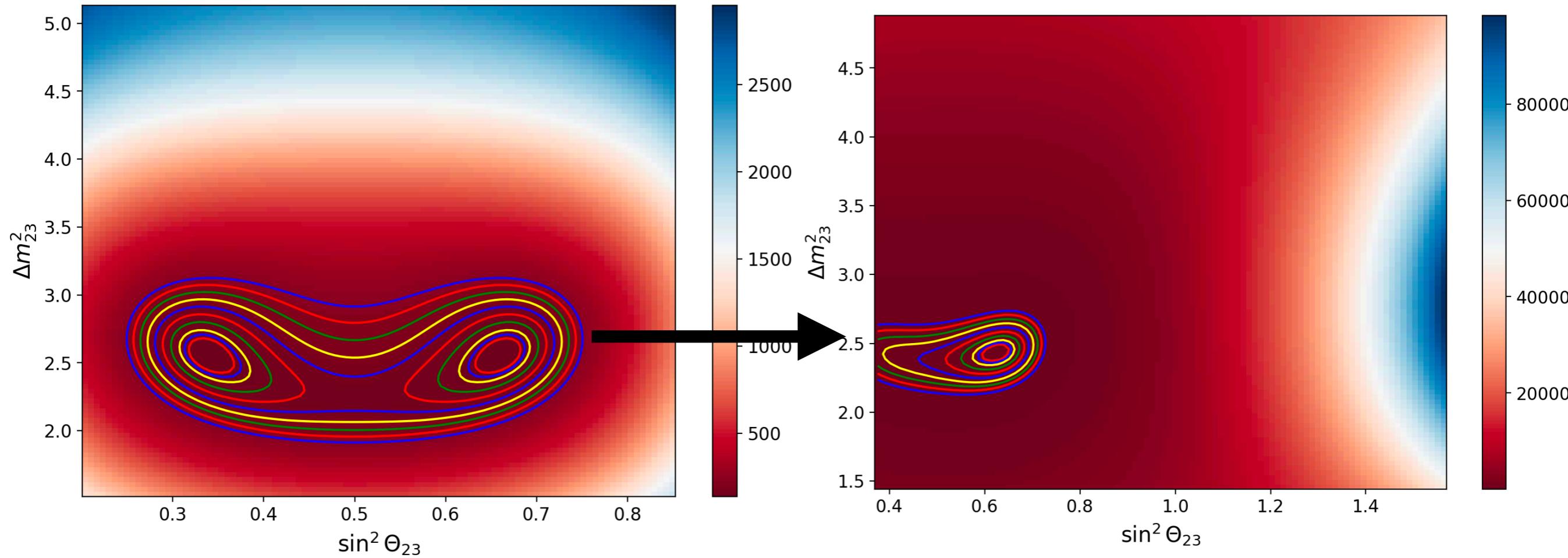


Frequentist View:
Make sure your model is
complete

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

Float this parameter

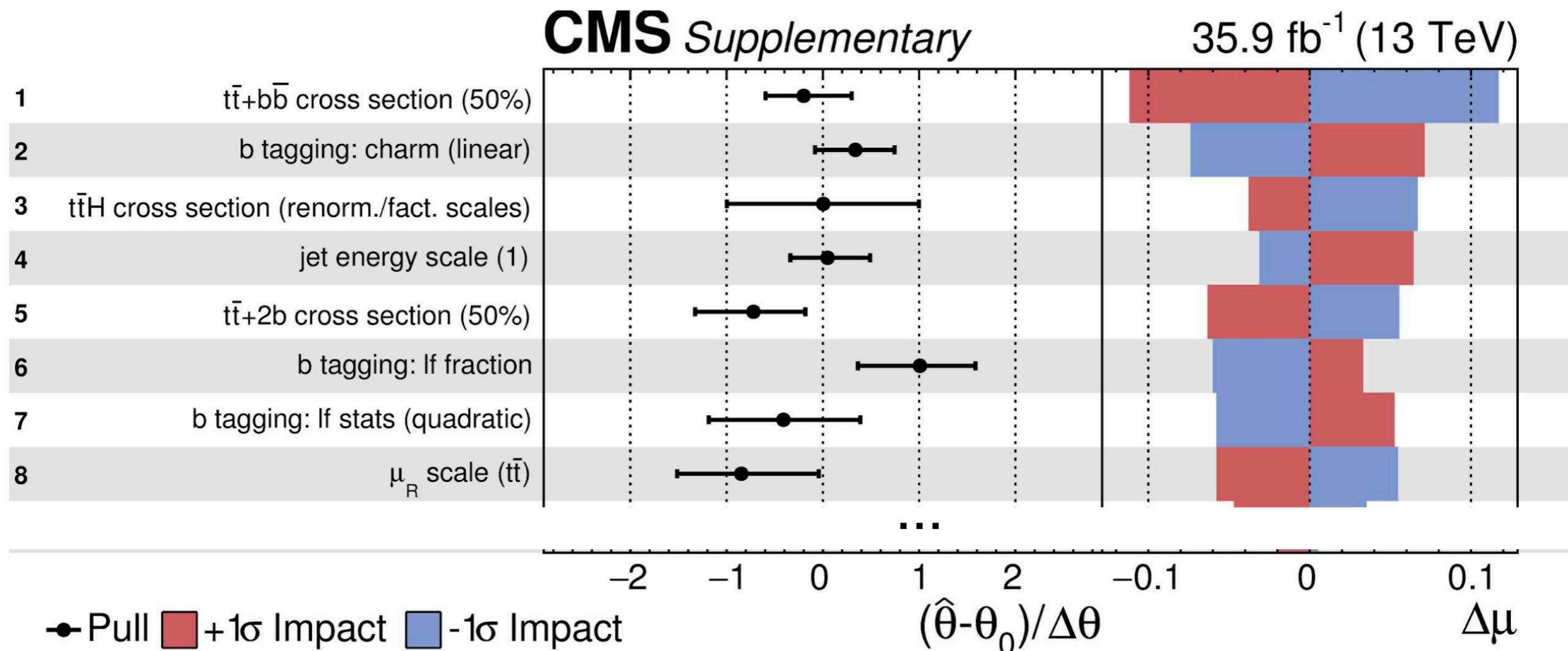
Nuisance



External Nuisances

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

Nuisance

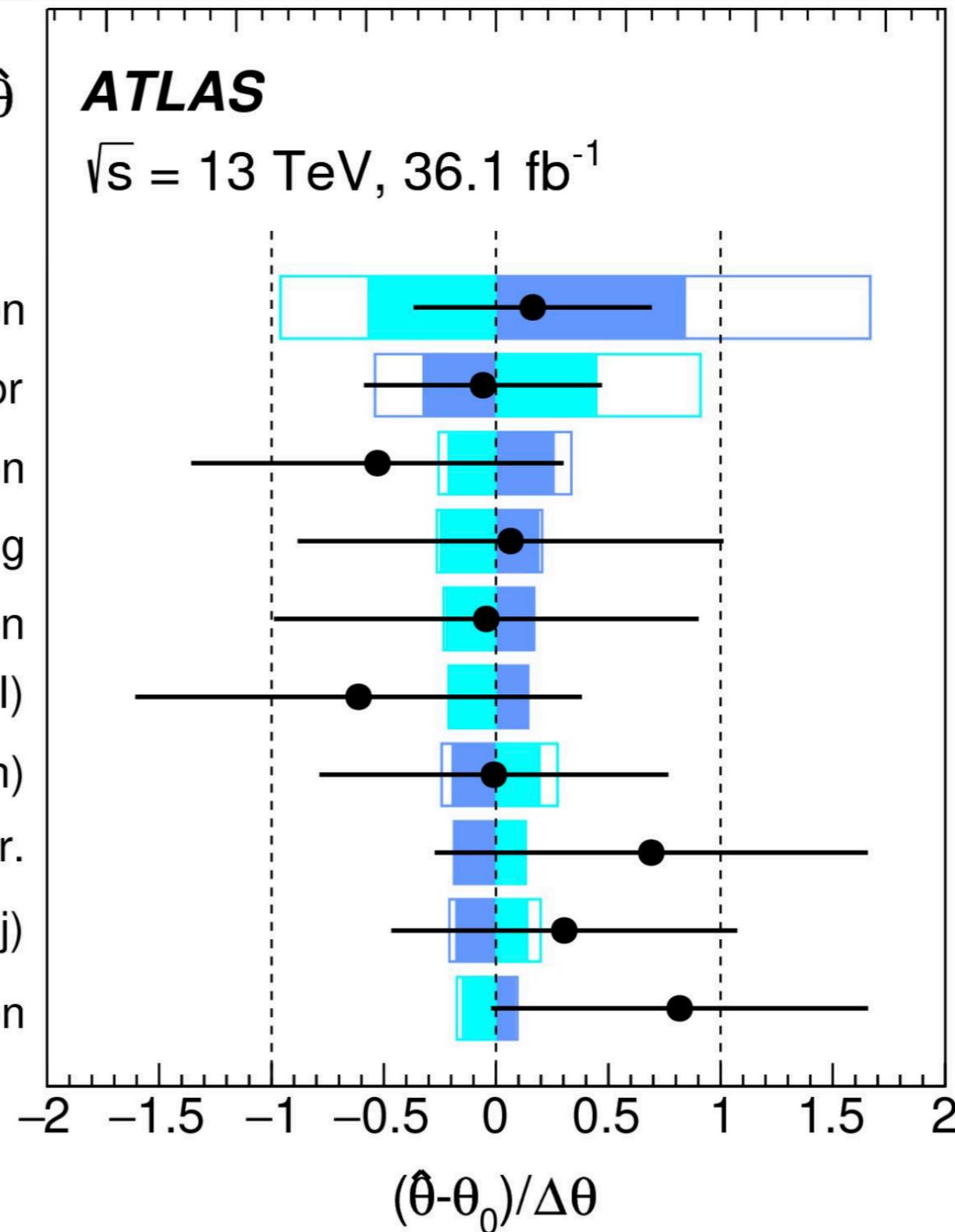


Nuisance

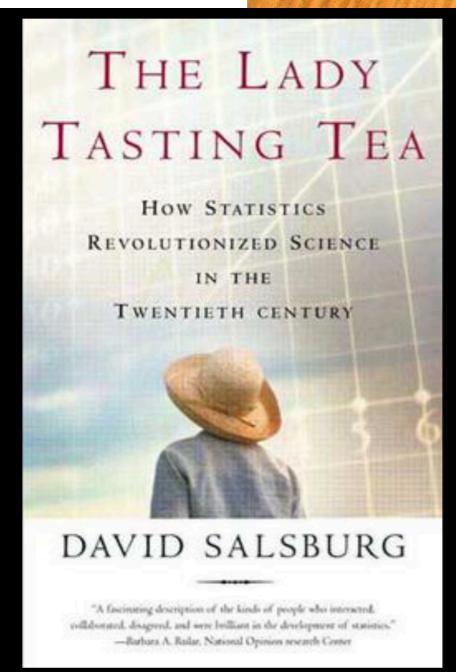
Post-fit impact on μ :

- █ $\theta = \hat{\theta} + \Delta\hat{\theta}$
- █ $\theta = \hat{\theta} - \Delta\hat{\theta}$
- Nuis. Param. Pull

$t\bar{t}$ +jets PS and hadronization
 $t\bar{t}$ +jets NLO generator
 $t\bar{t}+\geq 1c$ normalization
 $t\bar{t}+\geq 1b$ NLO reweighting
 $t\bar{t}+\geq 1b$ MPI normalization
 b -tagging efficiency (NP I)
JES (flavor composition)
 $t\bar{t}$ +light / $\geq 1c$ NNLO top- p_T corr.
 W/Z +jets normalization (9j)
 $t\bar{t}+\geq 1b$ normalization



Hypothesis Test

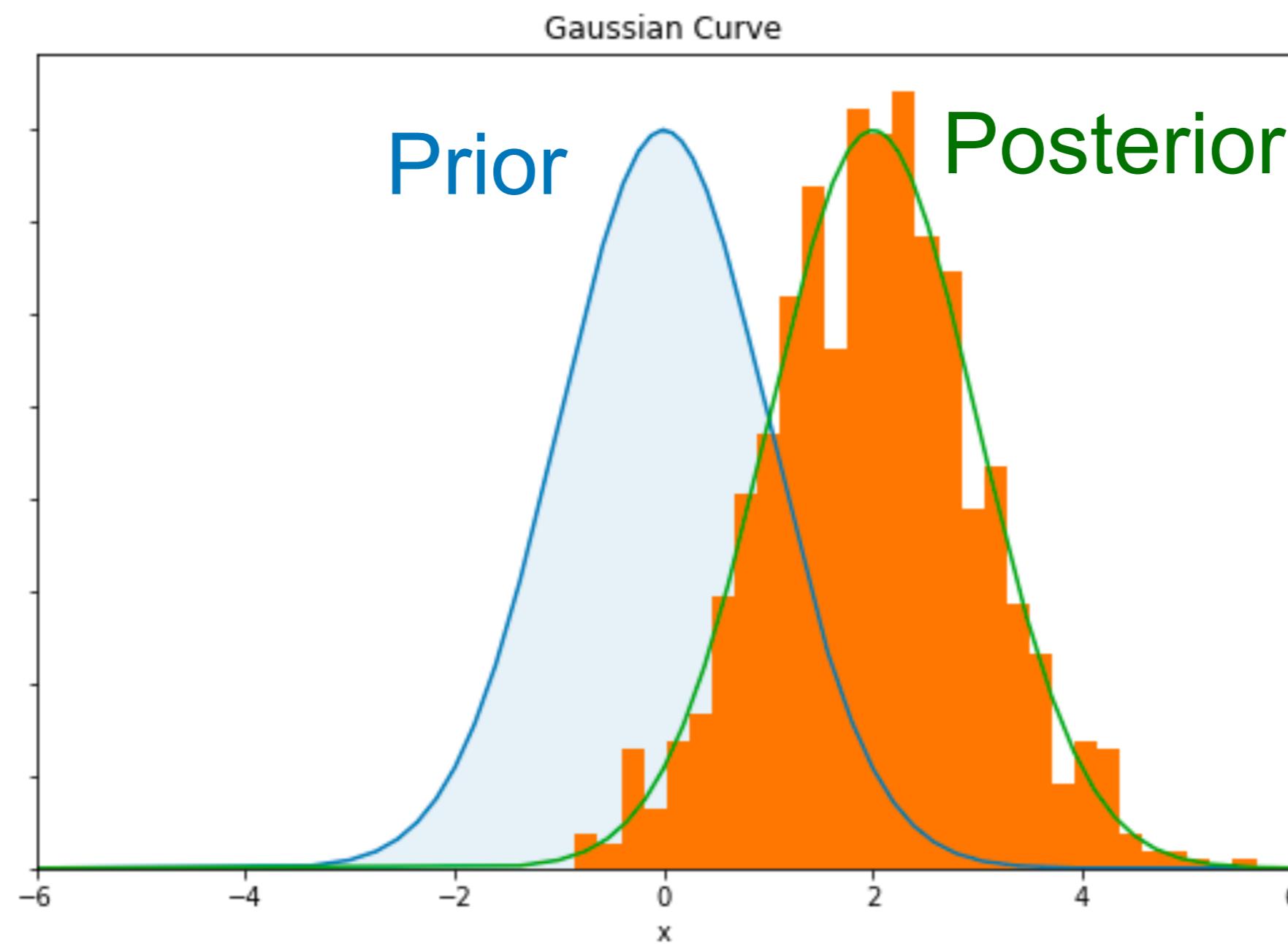


Hypothesis Test

- 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?**

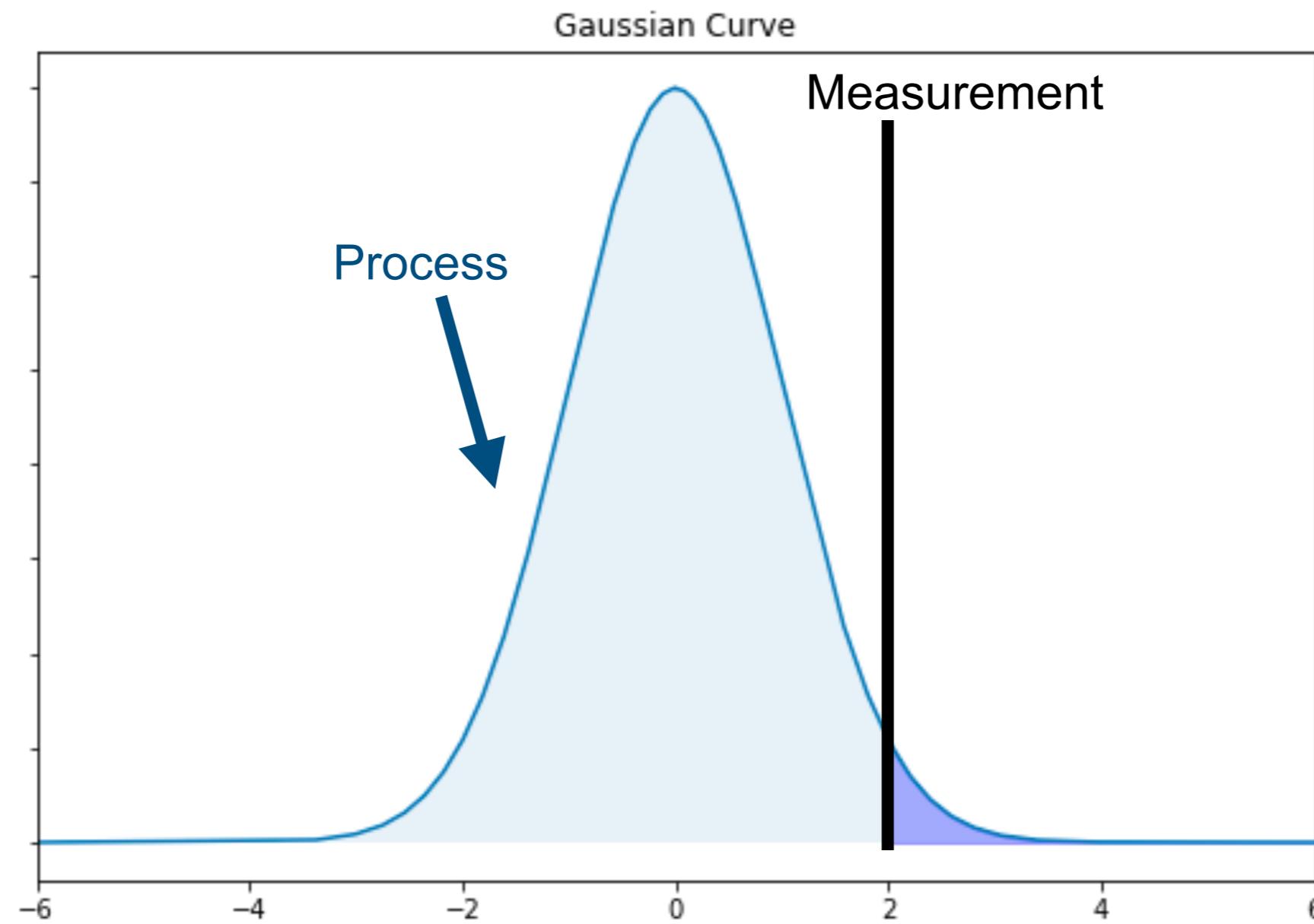
Lets Get Back to the³¹ original

- How do I tell if my measurements are consistent?



What is a measurement?

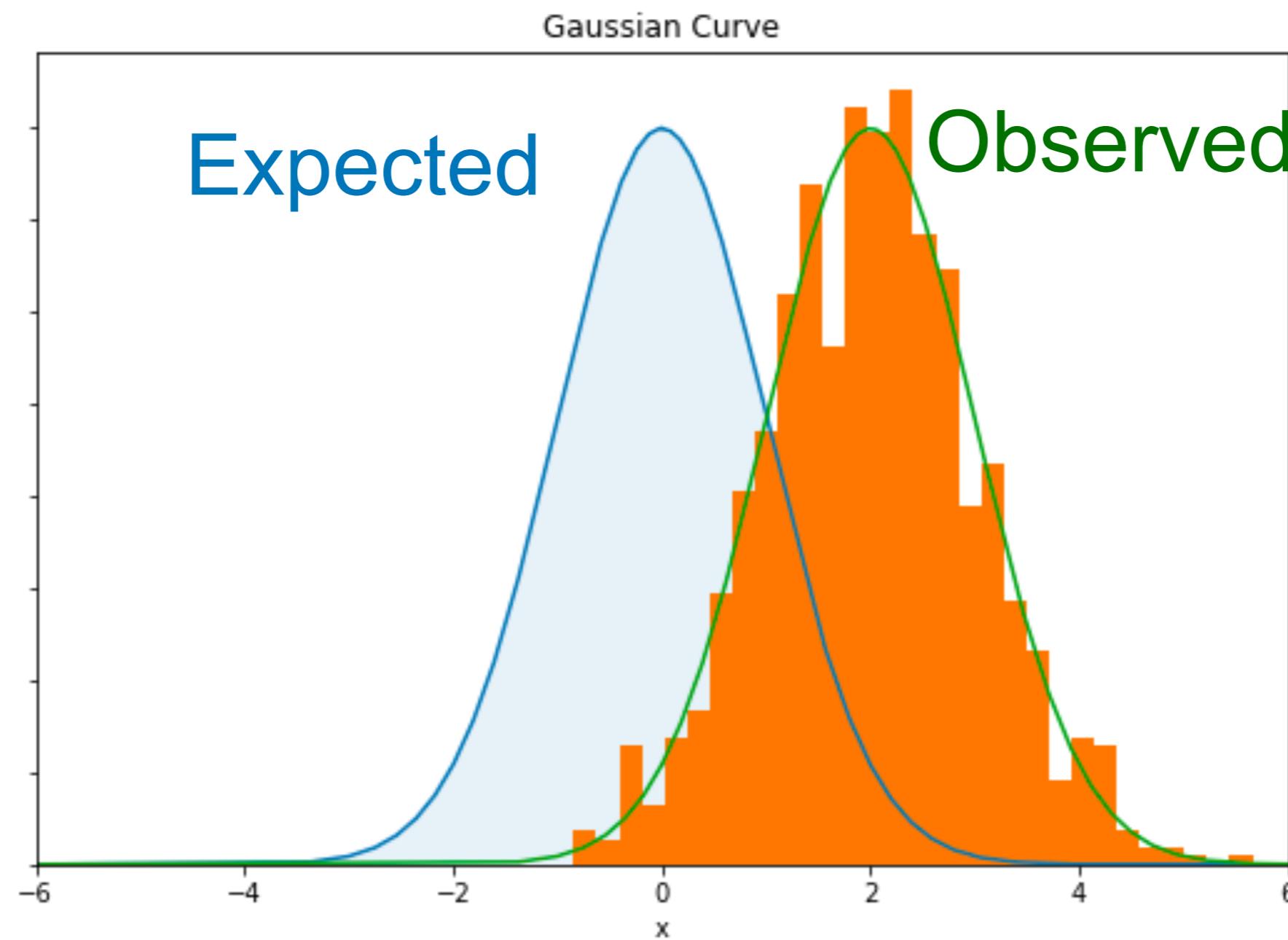
- We have assumed a model for this process
 - Example what is the direction of your eye reading slide?



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

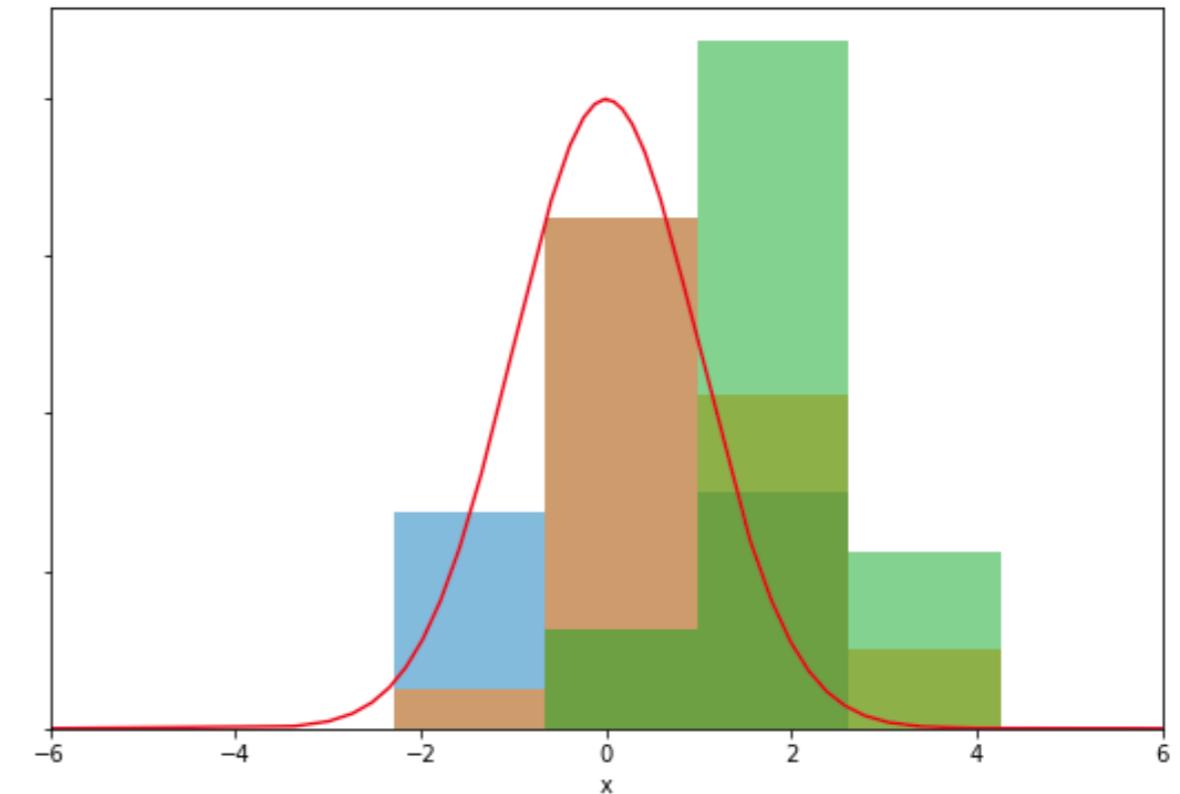
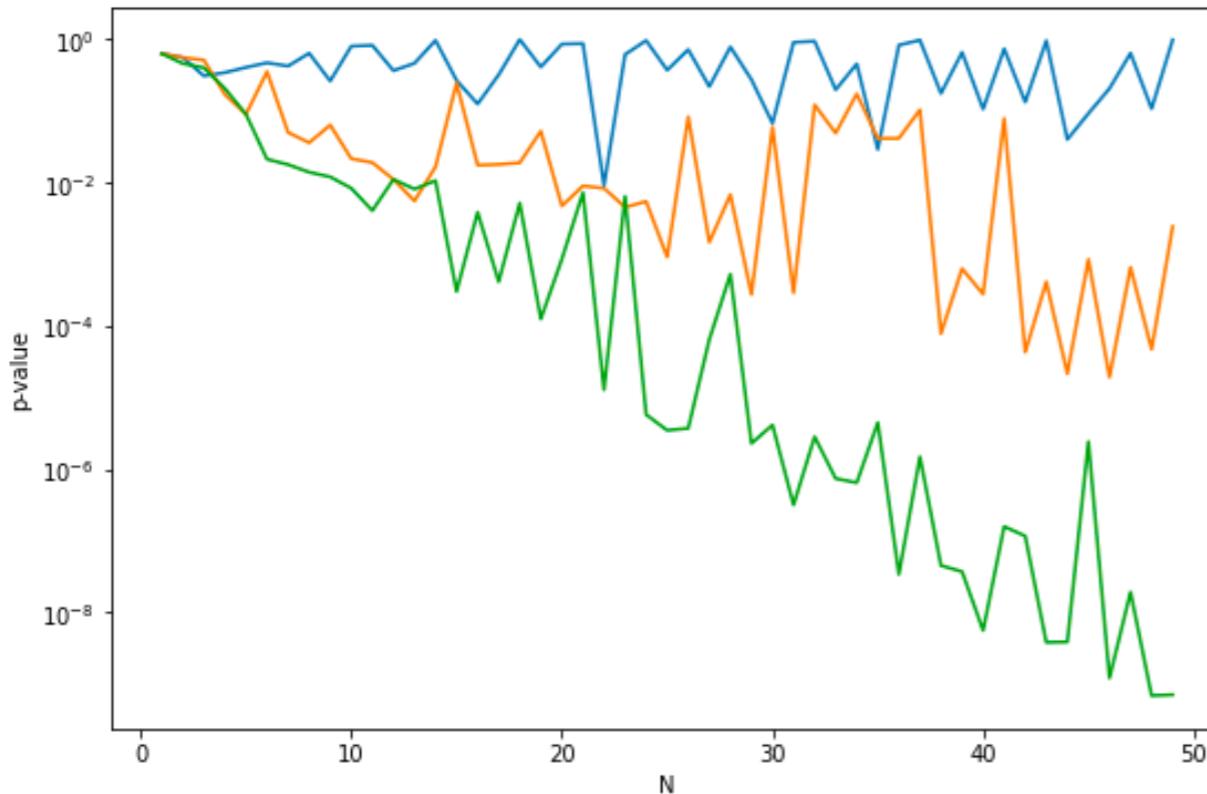
$E_i = Np_i$ = the expected (theoretical) count of type i , asserted by the null hypothesis that the fraction of type i in the population is p_i

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

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

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

↑
x

Supremum Function (ie max difference)

