

CLASS 10

ANALYTICAL OLS, LIN-REG WITH BAYESIAN INFERENCE

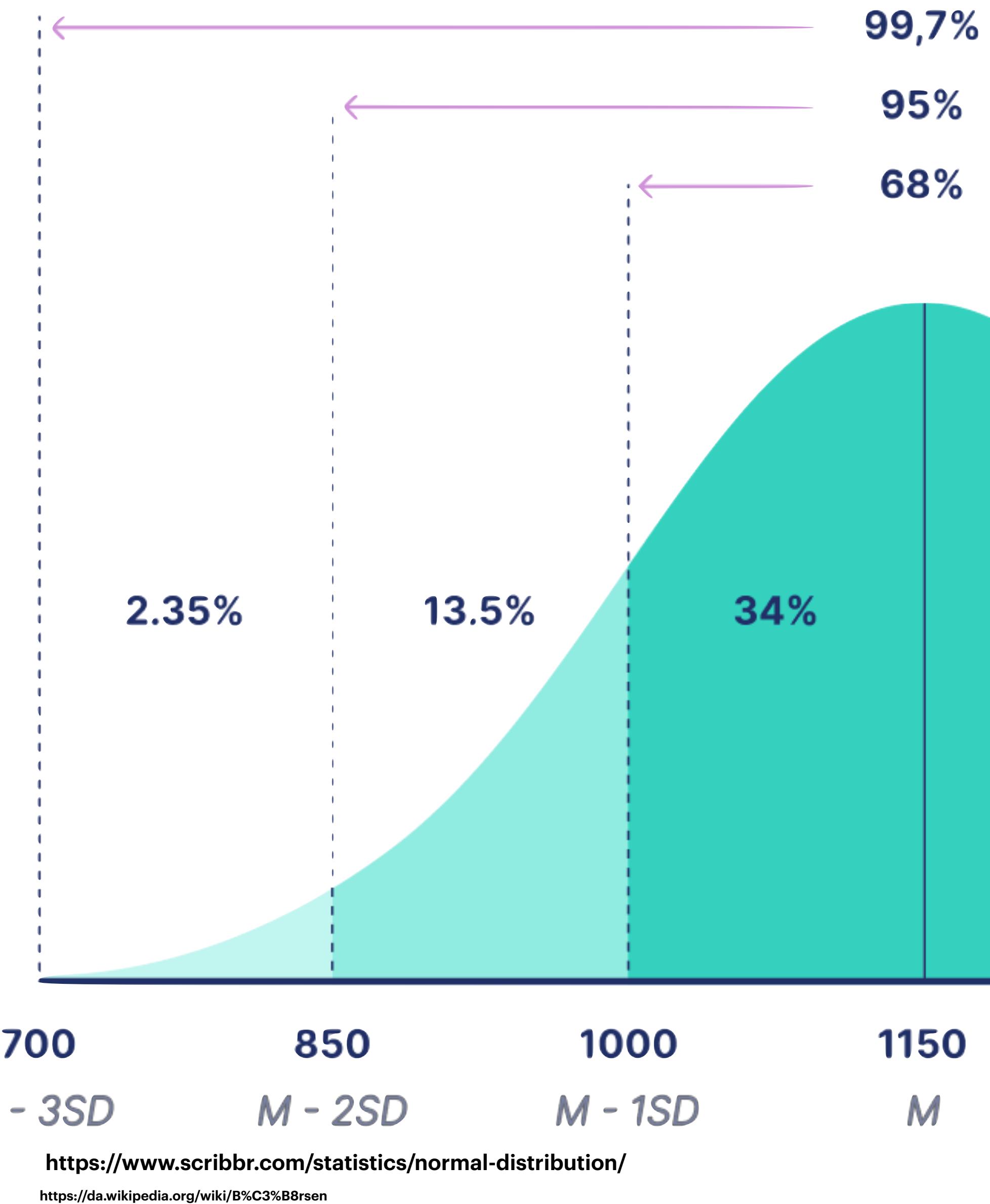


Albert Camus

Skud ud 

AGENDA

- Anything to recap?
- Linear models and Bayesian inference
 - lm() vs. stan_glm()
- Exercises  + portfolio work
- More recap if time



LINEAR MODEL ON MATRIX FORM

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$\downarrow \quad \downarrow \quad \downarrow$

$$Y = X\beta + \varepsilon$$

Try to compute the vector and matrix multiplications and additions as we have learned in previous weeks if this is hard to grasp ;)



WHAT IS BAYESIAN INFERENCE?

Something with PRIORS??? What are priors

We use Bayes' theorem to update prob of a hypo, as more evidence becomes available

WHAT IS BAYESIAN INFERENCE?

Different from classical stats in that:

Classical stats: Probabilities are associated with the data ↗

Bayesian stats: Probabilities are associated with the hypotheses ?

Something with PRIORS??? What are **priors**

We use Bayes' theorem to update prob of a hypo, as more evidence becomes available

Prior distribution: What do you know about the parameters before we see the data?

WHAT IS BAYESIAN INFERENCE?

Different from classical stats in that:

Classical stats: Probabilities are associated with the data ↗

Bayesian stats: Probabilities are associated with the hypotheses ?

EXAMPLE OF BAYESIAN INFERENCE: EXPLAINING WHAT PRIORS/POSTERIORS ARE USING HEIGHTS

If it's hard to grasp (this is difficult stuff) ➤❗️➤



Who We Are What We Do Industries



Analytics, Data Science

Bayesian vs. Frequentist Methodologies Explained in Five Minutes

<https://infotrust.com/articles/bayesian-vs-frequentist-methodologies-explained-in-five-minutes/>

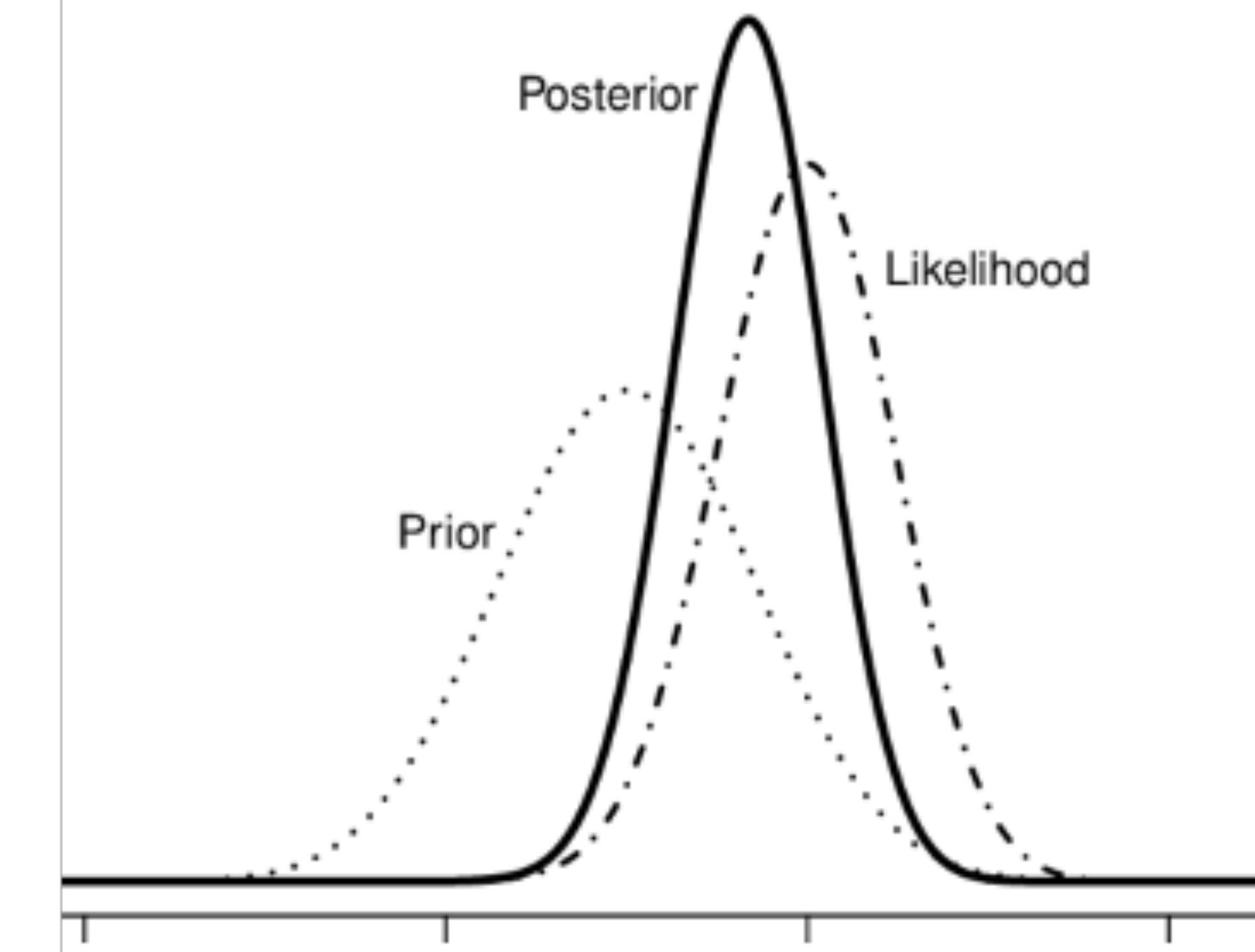
BAYESIAN INFERENCE

BAYES THEOREM

$$p(\theta|Data) = \frac{p(Data|\theta) p(\theta)}{p(Data)}$$

Diagram illustrating the components of Bayes' Theorem:

- Posterior: The resulting distribution after combining the likelihood and prior.
- Likelihood: The probability of the observed data given a specific parameter value.
- Prior: The initial beliefs or knowledge about the parameter before seeing the data.
- Normalization: The factor used to scale the joint distribution so that it integrates to 1.



Remember theta is the general term for model parameters, which in the case of the linear model are the betas

BAYESIAN INFERENCE

Bayes' theorem ...

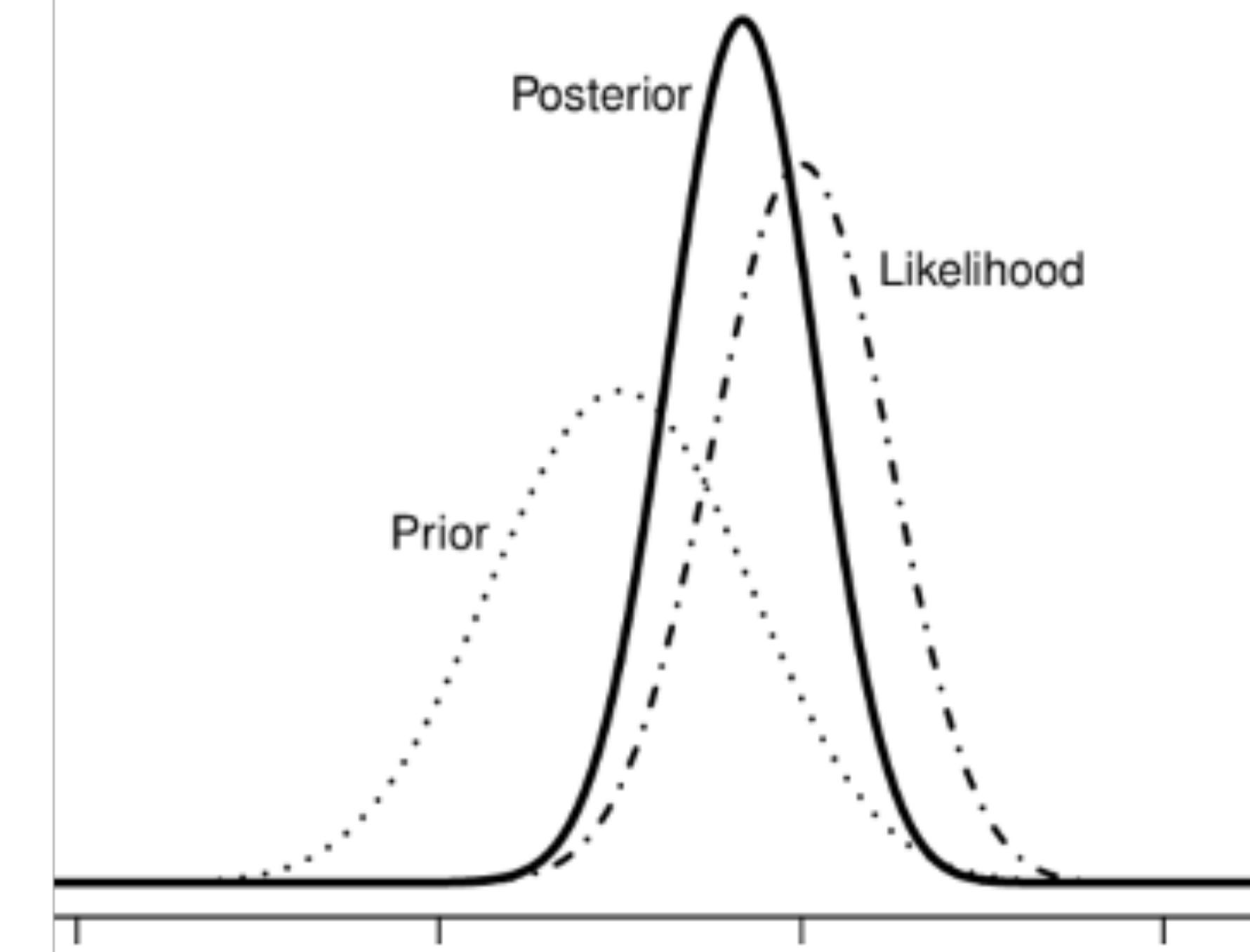
- 1) updates our beliefs about parameters (the **priors**)
- 2) with new info from the data (the **likelihood**)
- 3) to give us a revised belief (the **posterior**)

BAYES THEOREM

$$p(\theta|Data) = \frac{p(Data|\theta) p(\theta)}{p(Data)}$$

Diagram illustrating the components of Bayes' theorem:

- Posterior**: The resulting distribution after combining prior knowledge with new data.
- Likelihood**: The probability of observing the data given a specific parameter value.
- Prior**: The initial beliefs or knowledge about the parameter before seeing the data.
- Normalization**: The process of adjusting the joint probability to ensure it sums to 1.



Remember theta is the general term for model parameters, which in the case of the linear model are the betas

BAYESIAN INFERENCE

BAYES THEOREM

$$p(\theta|Data) = \frac{p(Data|\theta) p(\theta)}{p(Data)}$$

Likelihood Prior
Posterior Normalization

Remember theta is the general term for model parameters, \

Bayes' theorem ...

- 1) updates our beliefs about parameters (the **priors**)
 - 2) with new info from the data (the **likelihood**)
 - 3) to give us a revised belief (the **posterior**)
- stan_glm() does this 3 step programme ^^^^

in contrast, lm() just fits lin-reg models using **what?**

(hint: 

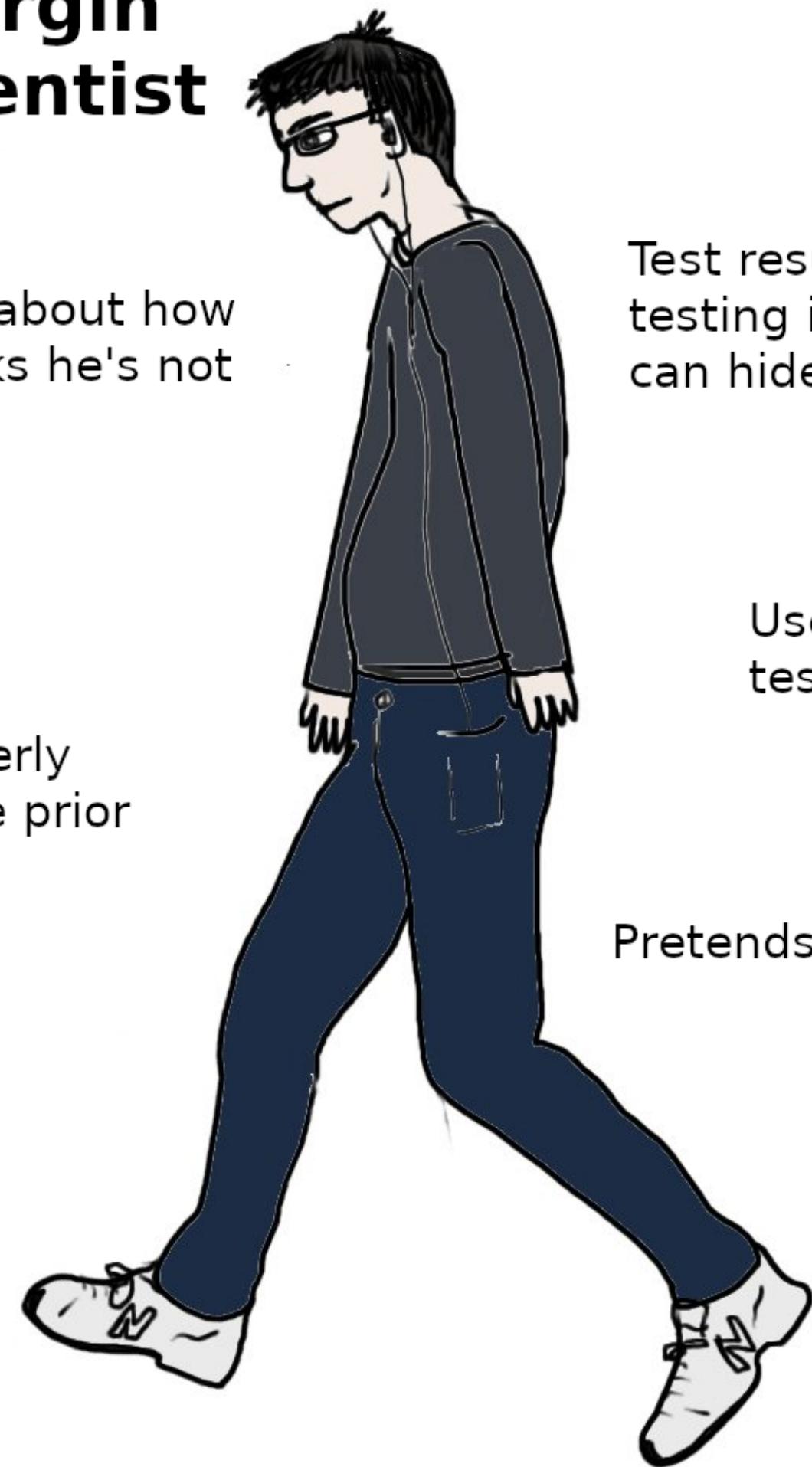
Why might **this (bayesian inf)** be better than?

- Incorporating prior knowledge allows building on previous work
- Probabilistic interpretation
- Many other reasons...



The virgin frequentist

Can only talk about how much he thinks he's not wrong



Can't properly incorporate prior research

Test results depend on testing intentions which he can hide

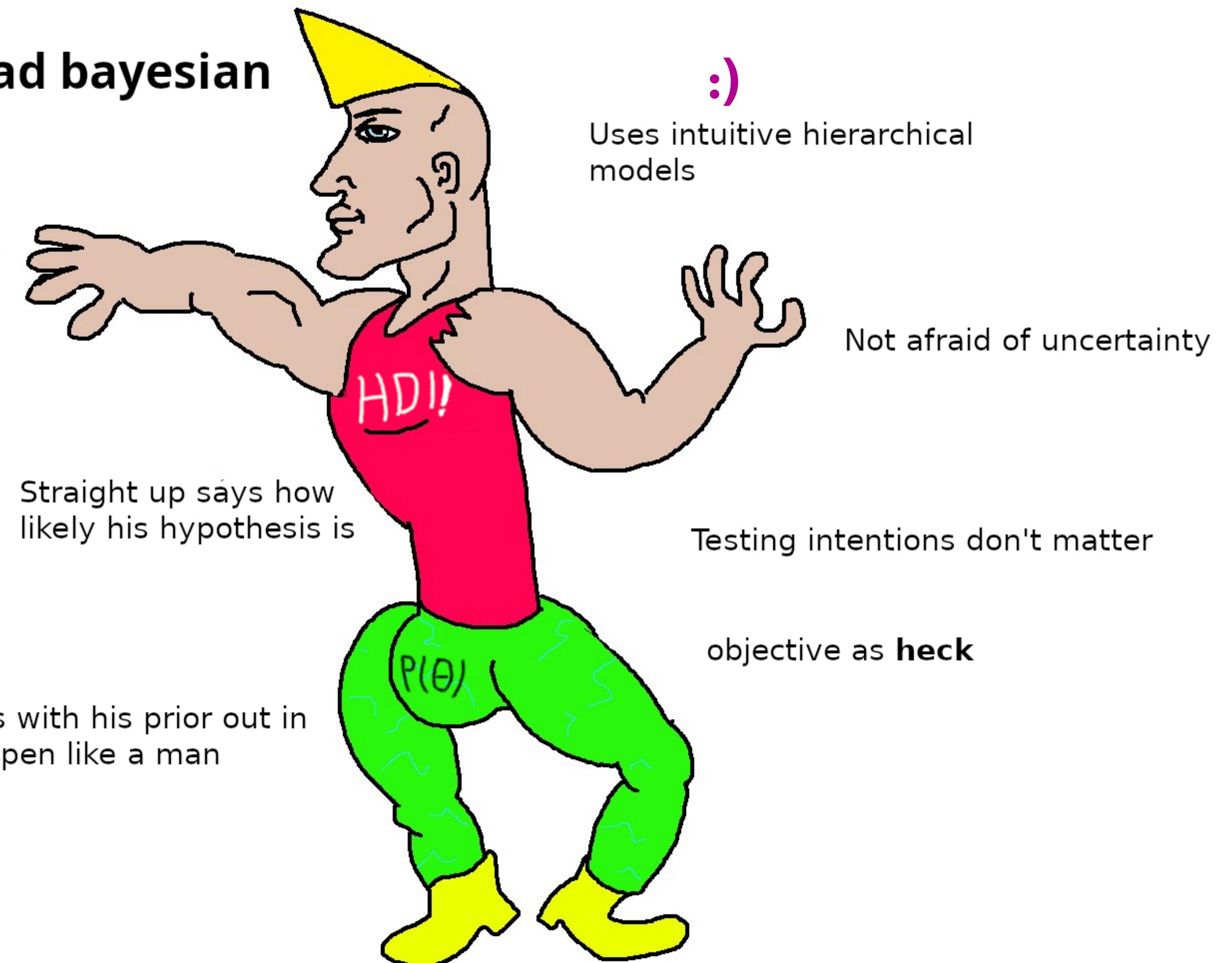
Uses unintuitive tests he's memorized

Pretends he's objective

The chad bayesian

Straight up says how likely his hypothesis is

Deals with his prior out in the open like a man



:)
Uses intuitive hierarchical models

Not afraid of uncertainty

Testing intentions don't matter

objective as **heck**

LM() VS STAN_GLM()

LM

Uses classical least squares and returns estimates and standard errors.

STAN_GLM

Performs bayesian inference and returns estimates, standard errors and posterior distributions.

Useful for visualizing uncertainty

Prior information can be incorporated

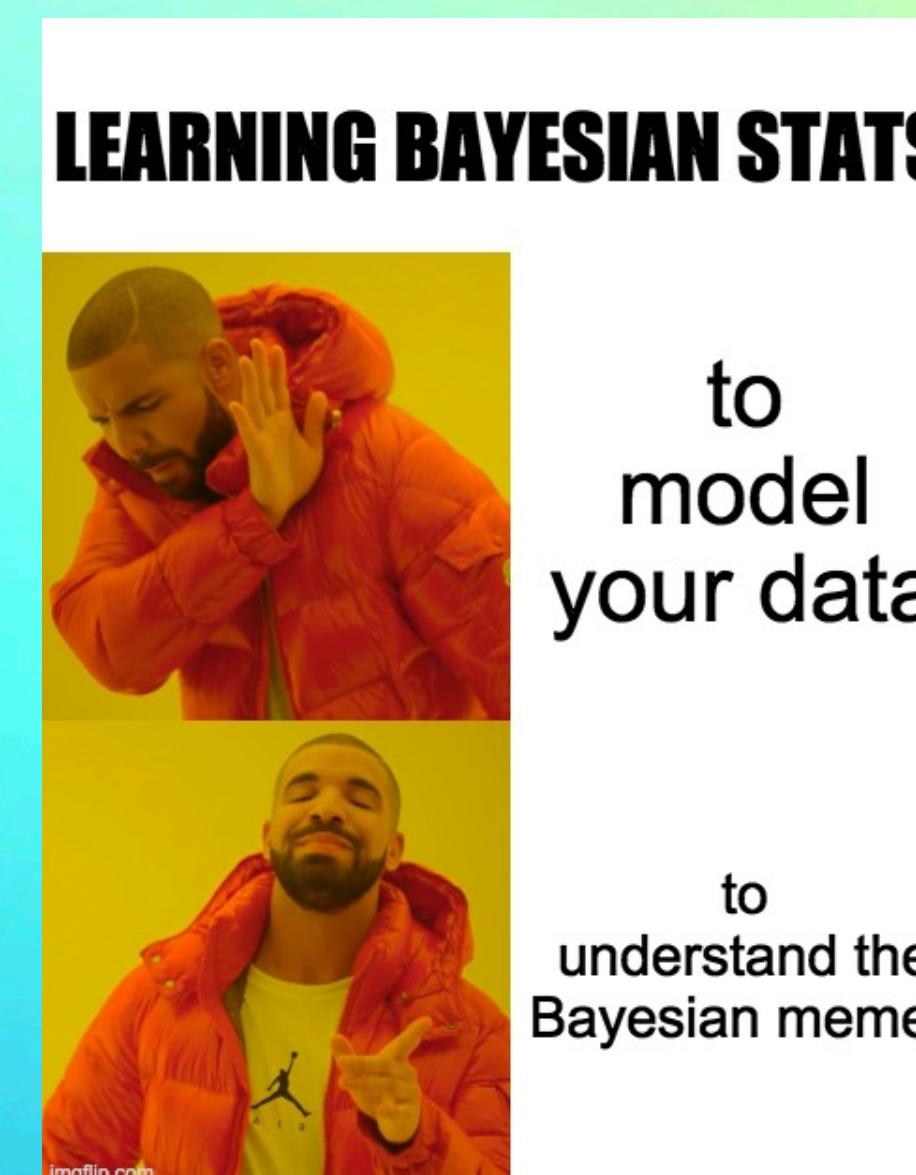
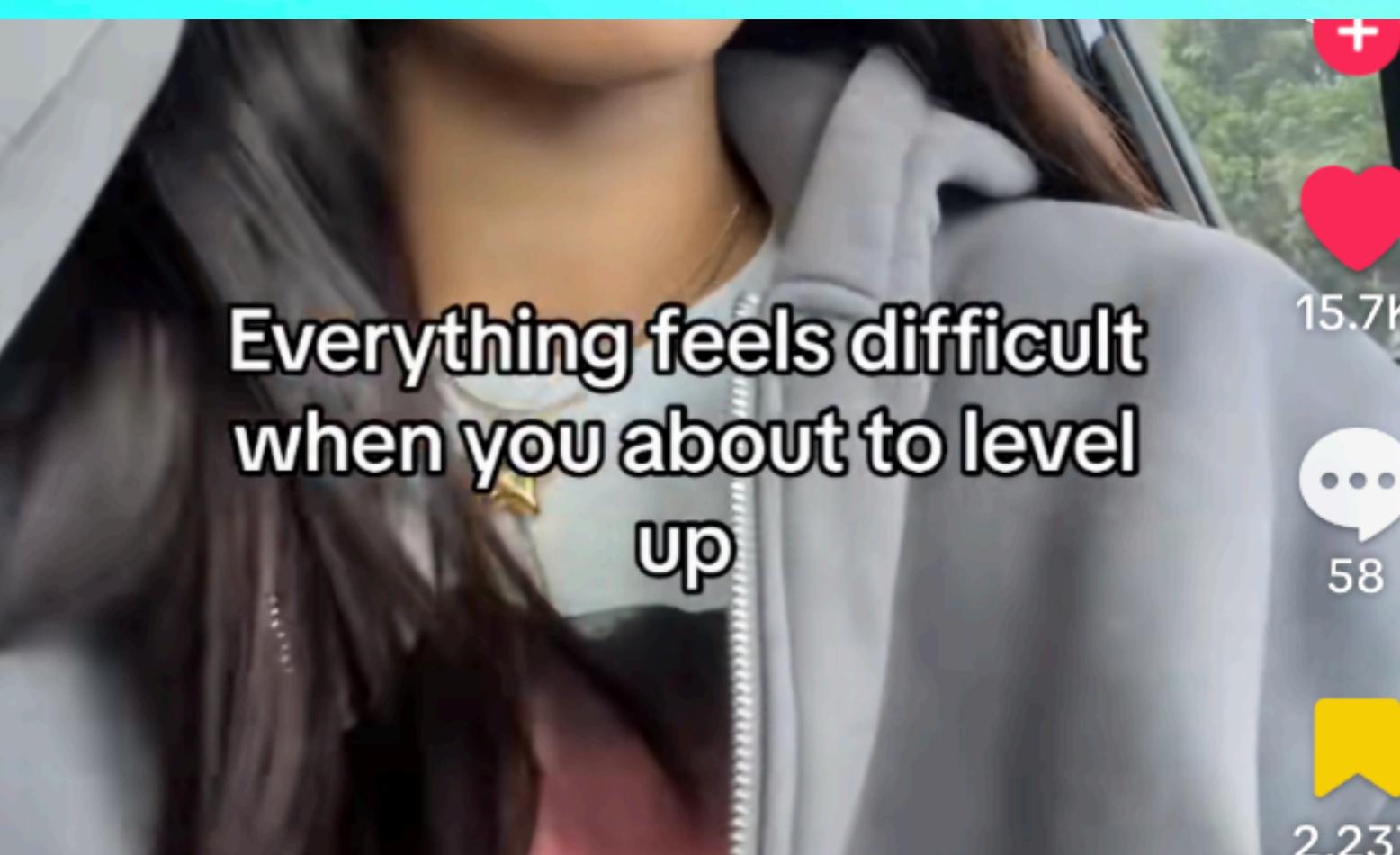
- For simple problems with no prior the LM and STAN_GLM will yield the same results

PRIOR VS POSTERIOR



EXERCISES

See .Rmd markdown for this class (10) on Git, and if extra time: portfolio work



Public Playlist

methods 22222222

what will the vibes be??

Pernille Brams and 8 others • 8 likes • 69 songs, 4 hr 10 min

#	Title	Artist	Duration
1.1	October	Alessia Cara	4:01
2	Parking Lot	Anderson .Paak	3:54
3	Evil Spider	BENEE	2:33
4	Awaken	Big Wild	3:03
5	Illusion	The Blue Van	3:35
6	Can I Call You Tonight?	Dayglow	4:39
7	Ode to a Conversation Stuck in Your ...	Del Water Gap	3:19
8	The Dress	Dijon	3:05
9	End of Beginning	Djo	2:39
10	She Wants My Money	Dominic Fike	2:14

Collab: <https://open.spotify.com/playlist/5UUiKD15vyFwymQ4qLur9V?si=82f6d92034de4285&pt=87474f98050a67a4599d7dfb9fa852d0>