

# While you are waiting...

---

- **socrative.com**, room number **SIMLANG2016**

# Simulating Language

## Lecture 13: Iterated Bayesian Learning, culture, innateness

---

Simon Kirby

[simon@ling.ed.ac.uk](mailto:simon@ling.ed.ac.uk)



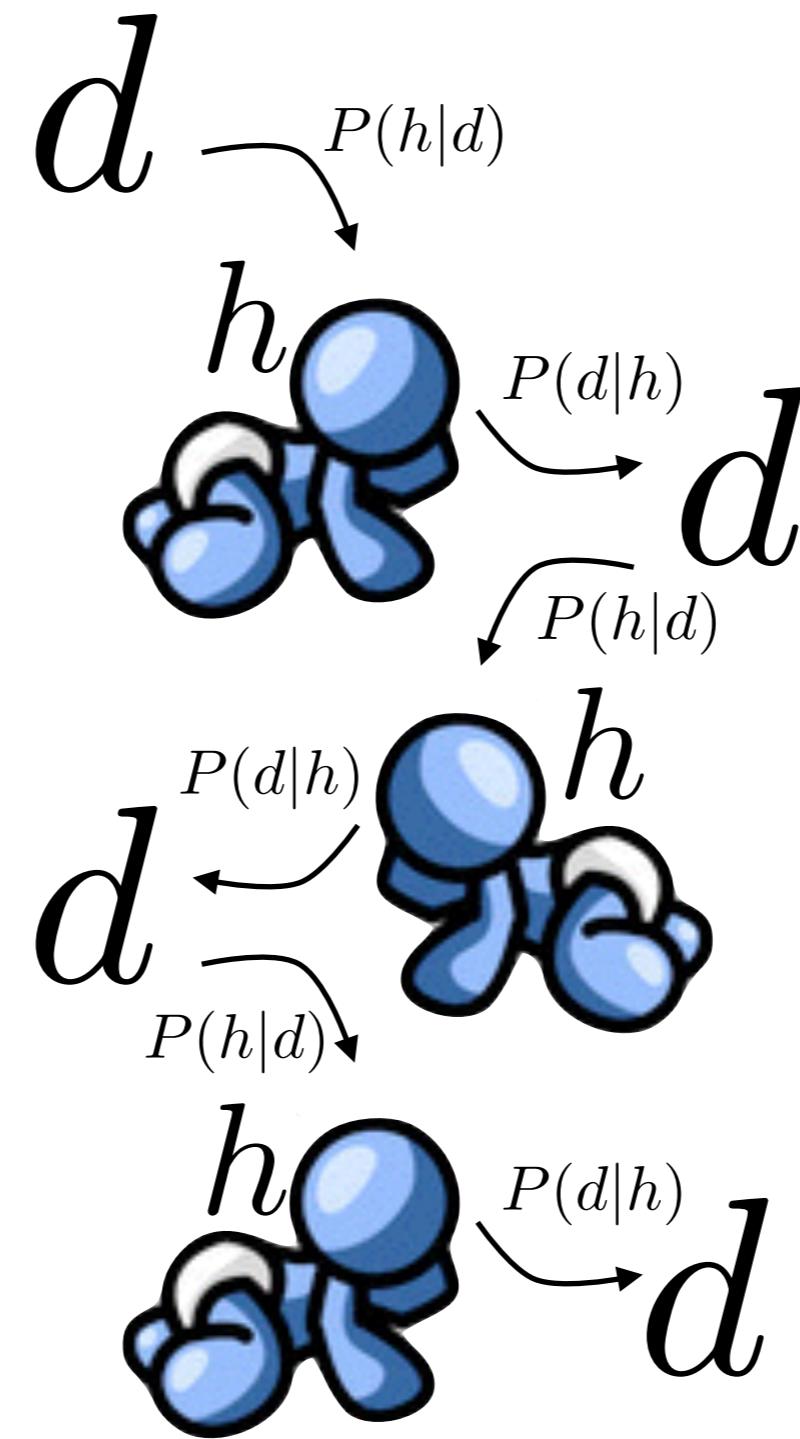
# Yesterday's lecture...

---

- We uncovered the importance of the *bottleneck* on cultural transmission
- It drives the evolution of structure because only structured languages can be stably transmitted through a bottleneck (without a bottleneck, language could stay holistic)
- This is a case of adaptation for learnability by a culturally evolving language

# A reminder: Iterated Bayesian Learning

---



# First results (Griffiths & Kalish 2007)

---

- Try out different models of language, different bottlenecks, different amounts of noise
- See how the process of cultural transmission takes the prior bias of the learner and gives rise to the actual resulting patterns of language
- What would you predict, based on the models you have seen so far?
- **The types of languages we see should:**
  - A. be completely unconstrained by the biases of language learners
  - B. reflect the biases of language learners, but in an interestingly complex way
  - C. directly reflect the biases of language learners

# First results (Griffiths & Kalish 2007)

---

- Try out different models of language, different bottlenecks, different amounts of noise
- See how the process of cultural transmission takes the prior bias of the learner and gives rise to the actual resulting patterns of language
- Their result:

Bottleneck does nothing

Noise does nothing

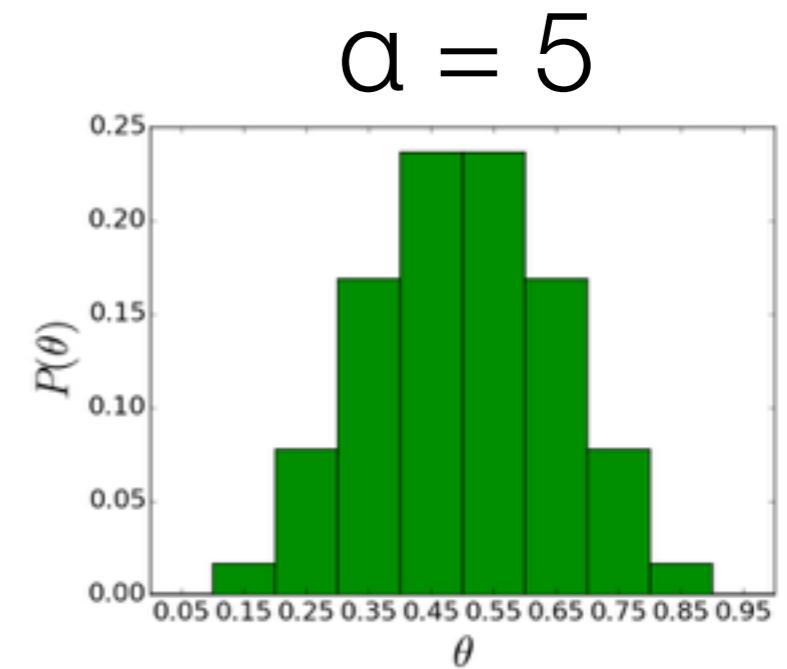
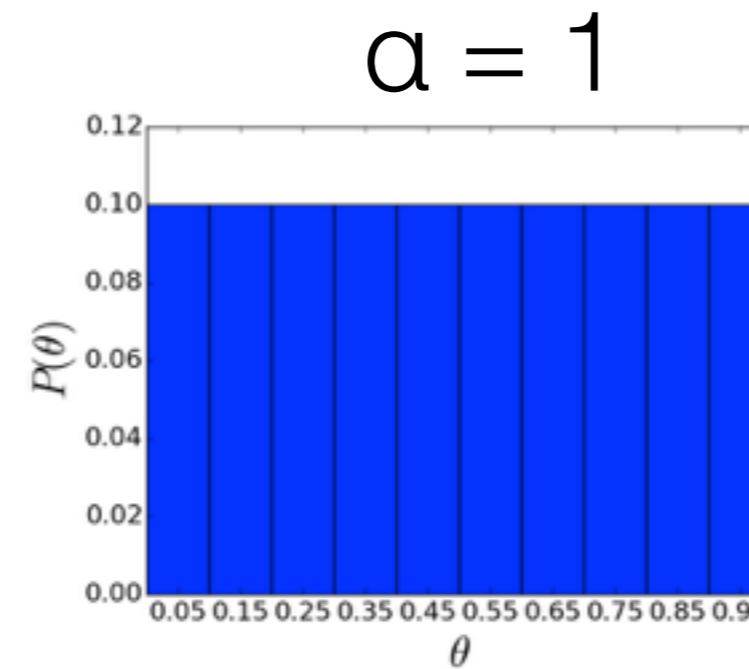
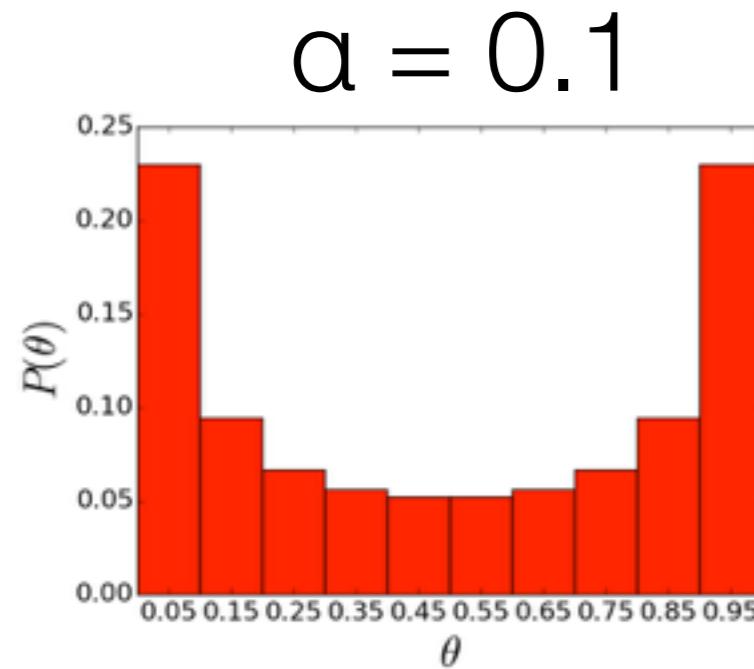
Details of language model do nothing

- Given enough time, the end result of cultural evolution always reflects the prior bias and nothing else

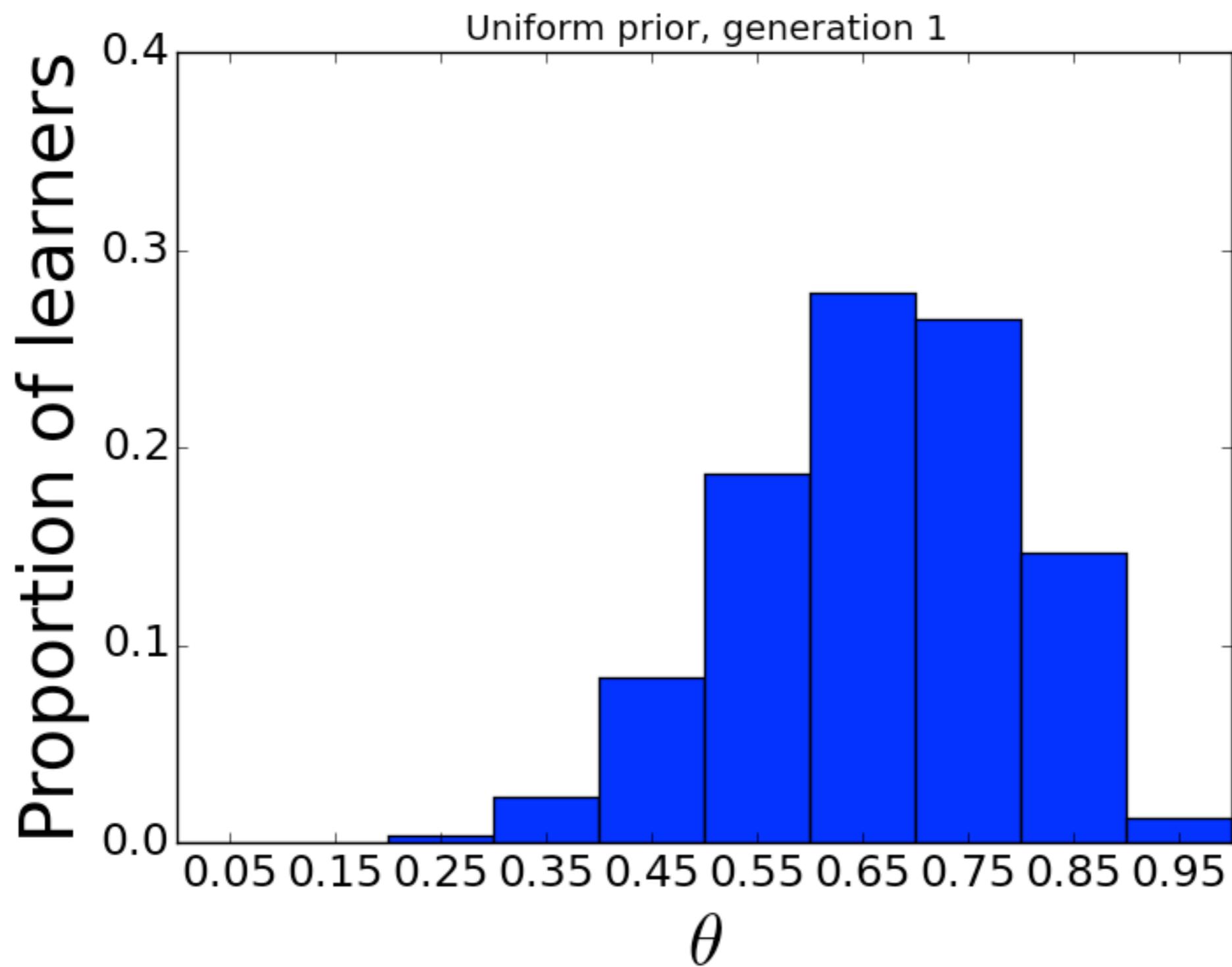
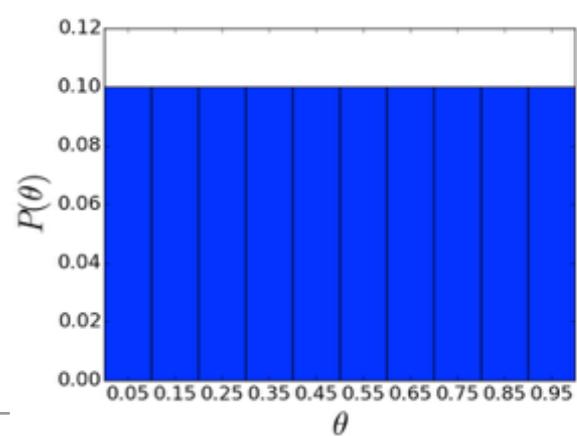
# You have already seen this result

---

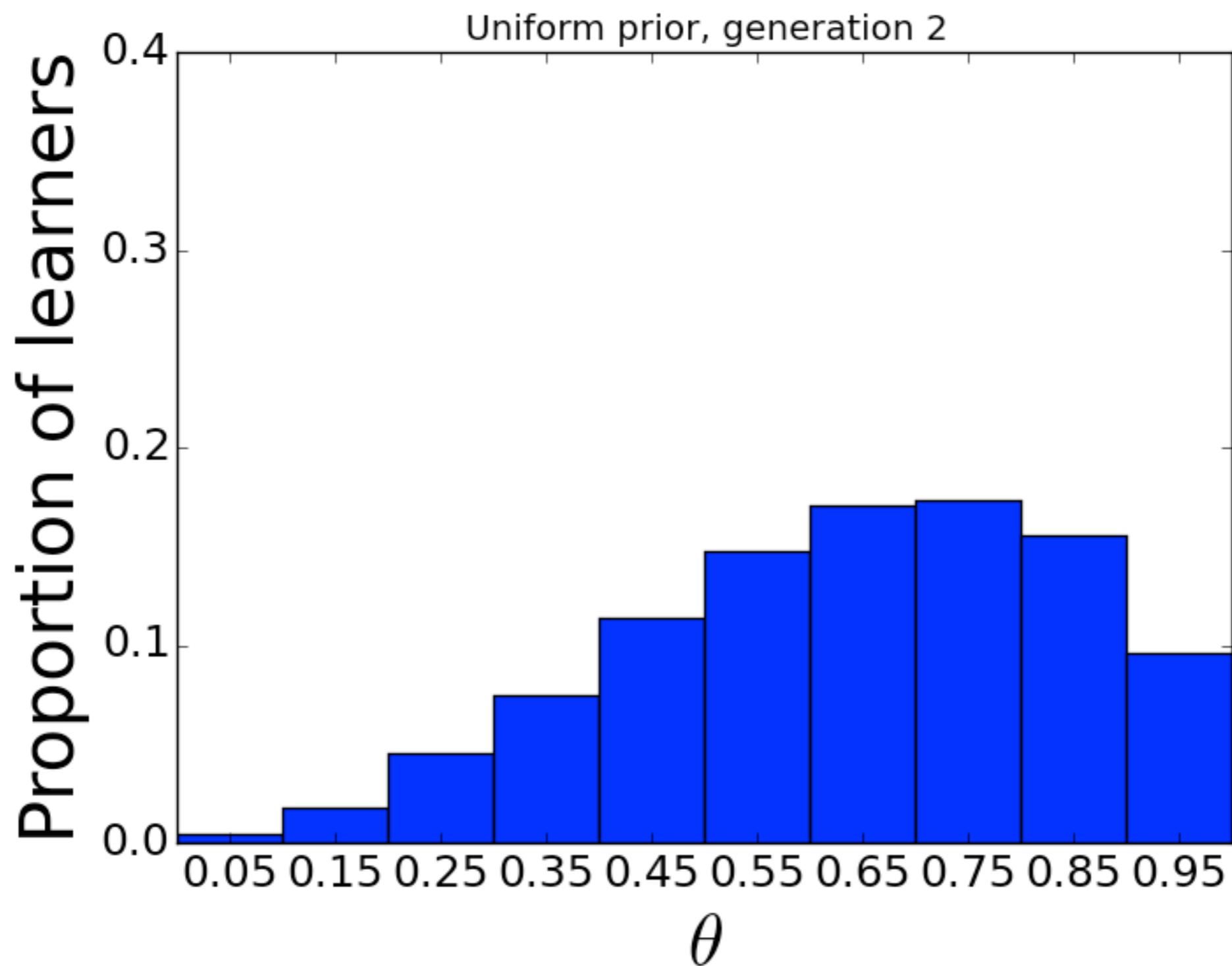
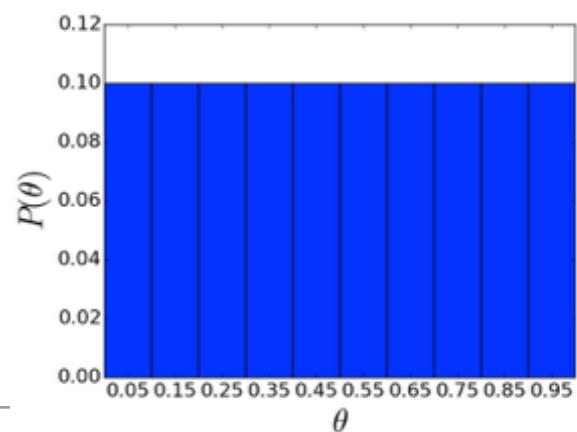
- Cast your mind back to lecture 10, beta-binomial model, learners estimating frequencies of two competing linguistic variants



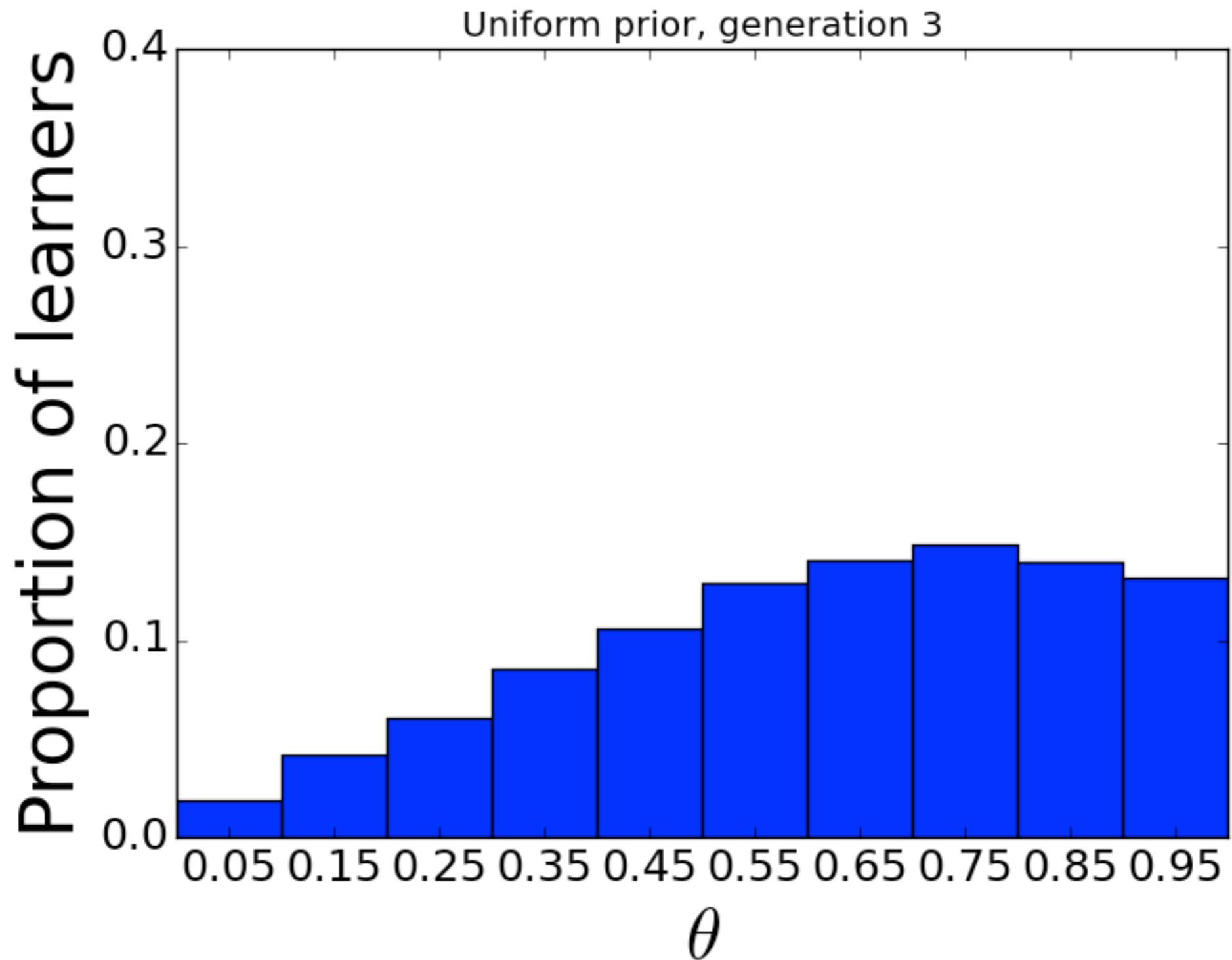
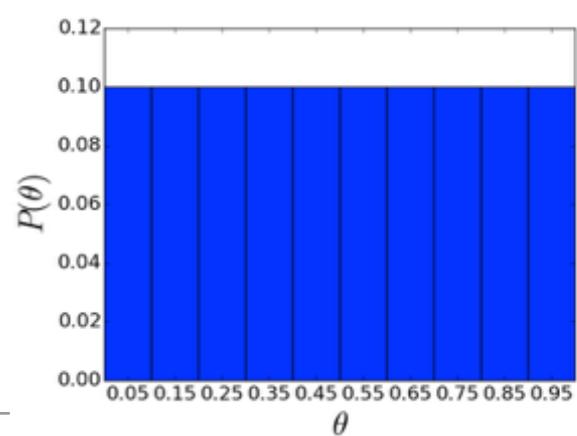
# Watching the prior reveal itself



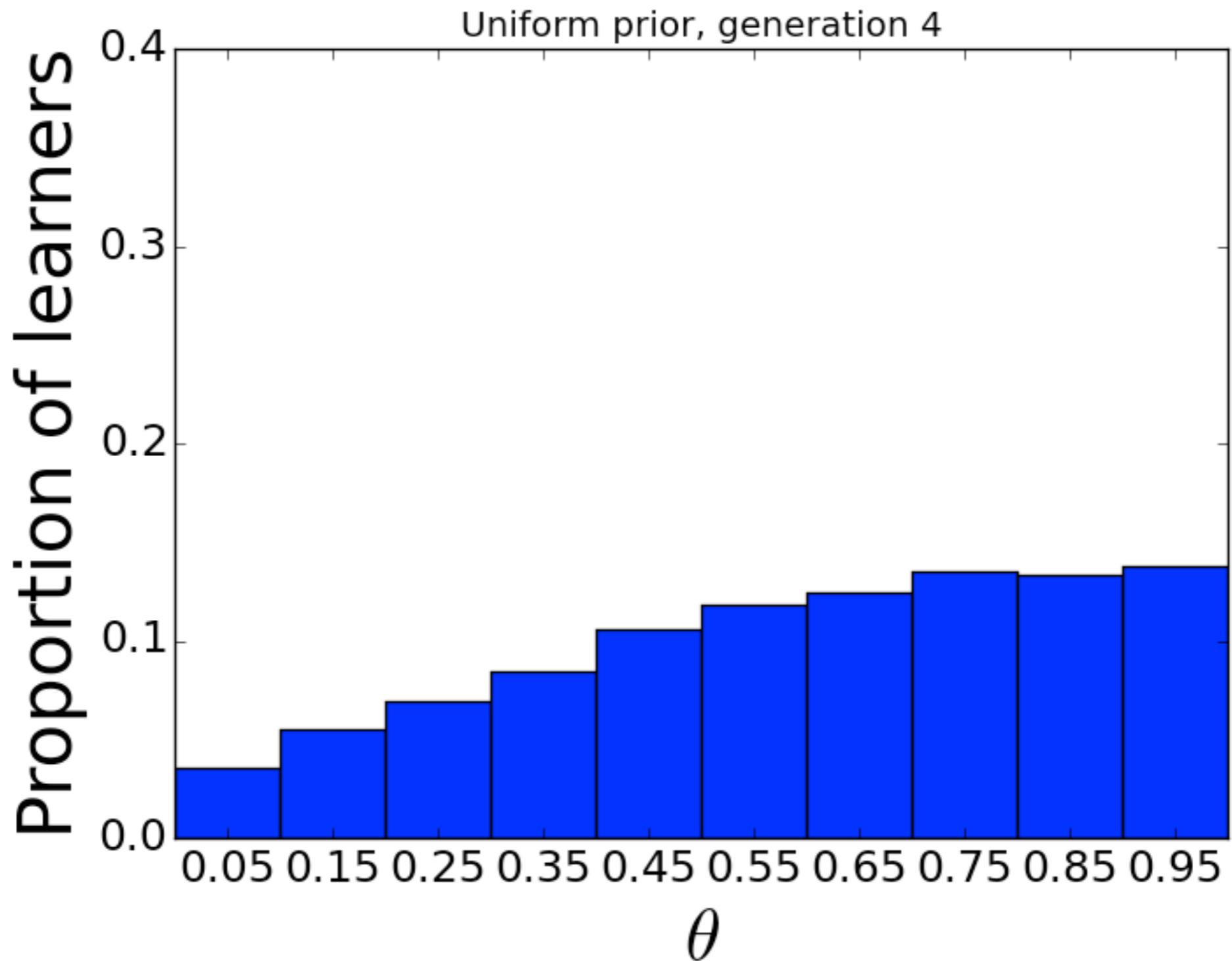
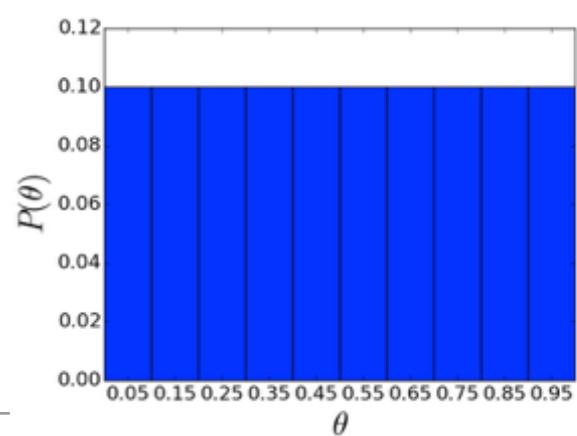
# Watching the prior reveal itself



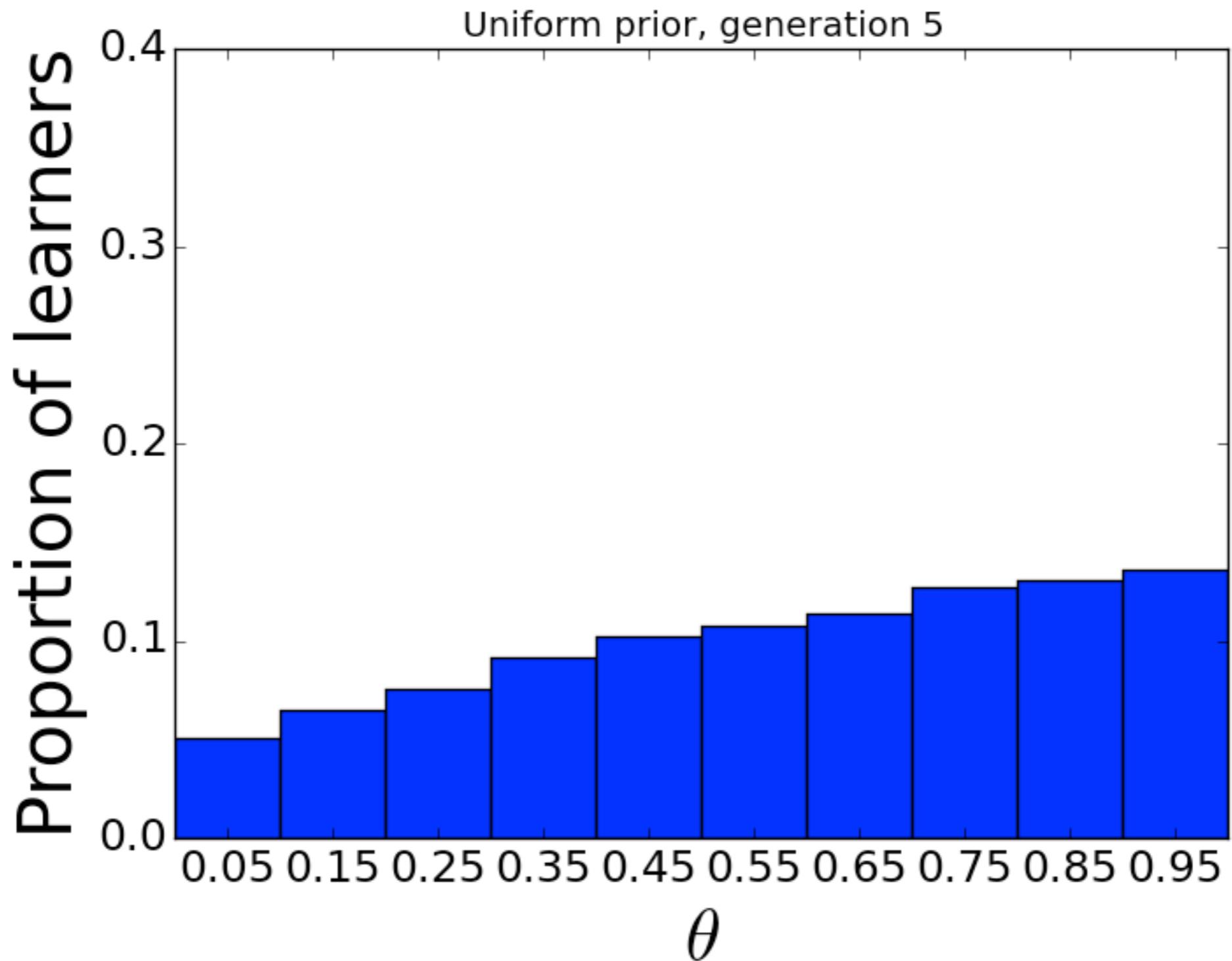
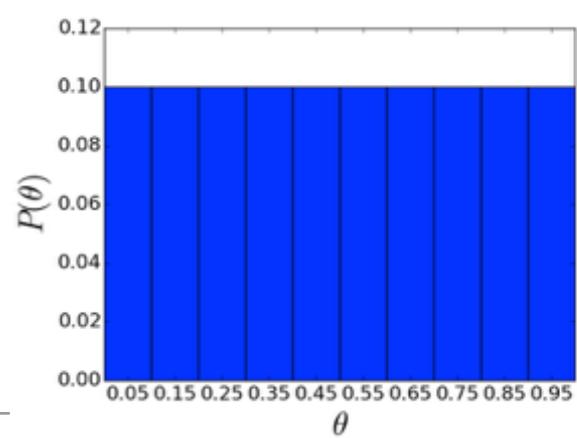
# Watching the prior reveal itself



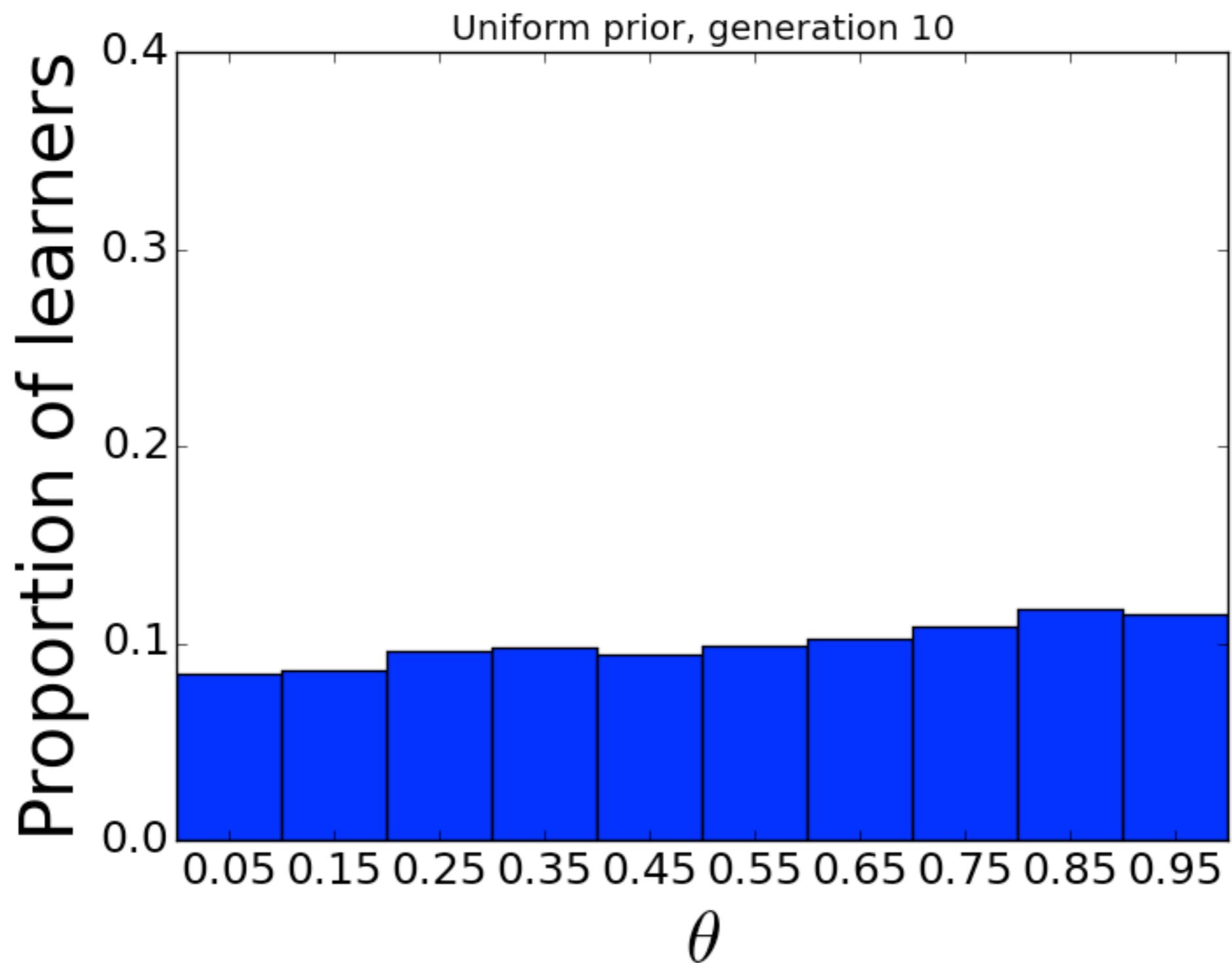
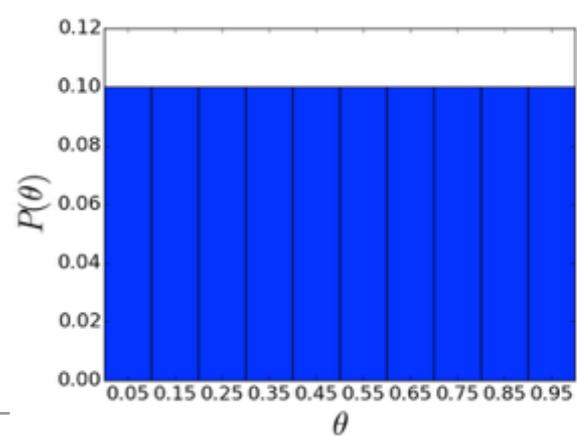
# Watching the prior reveal itself



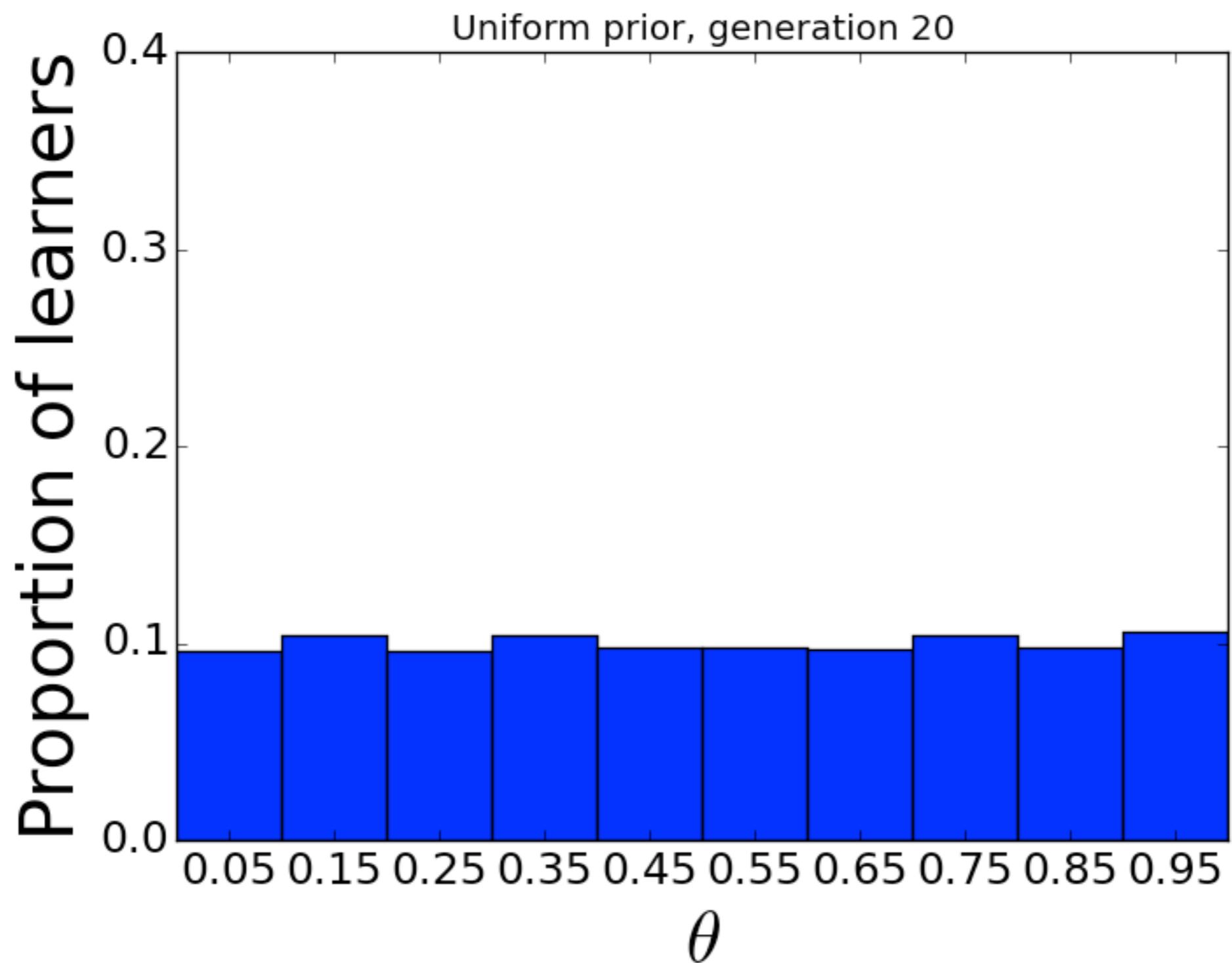
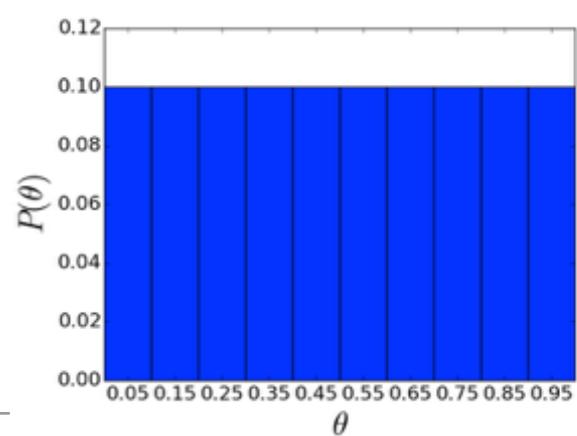
# Watching the prior reveal itself



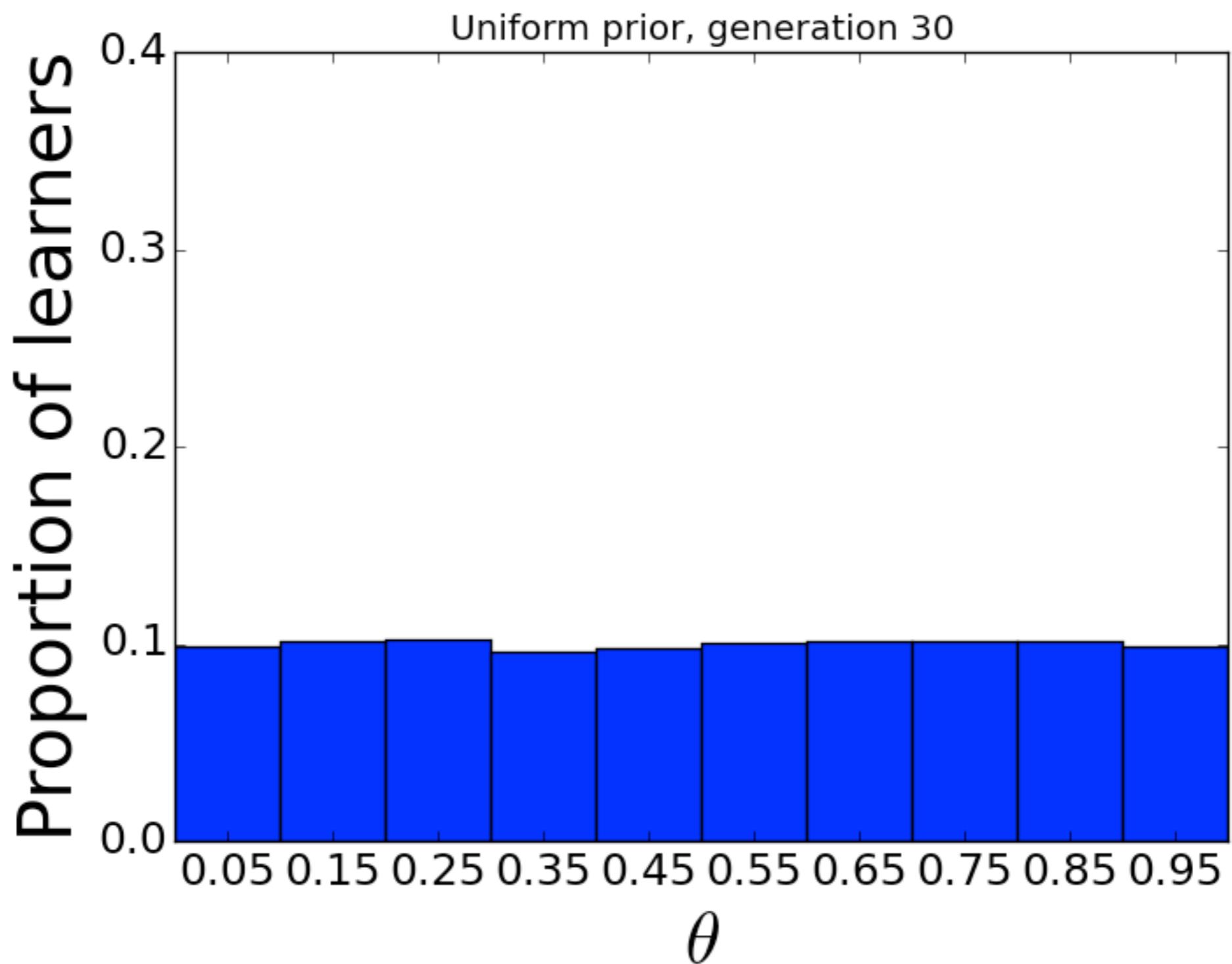
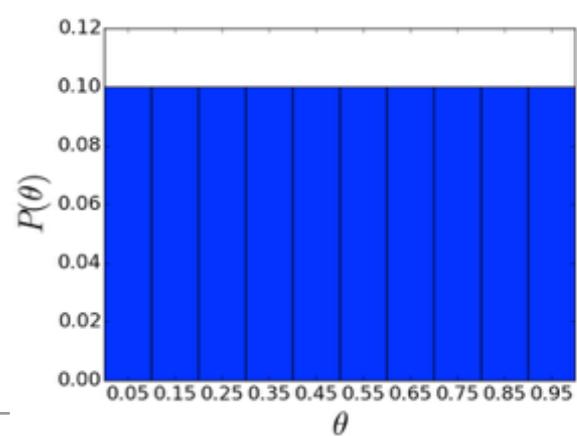
# Watching the prior reveal itself



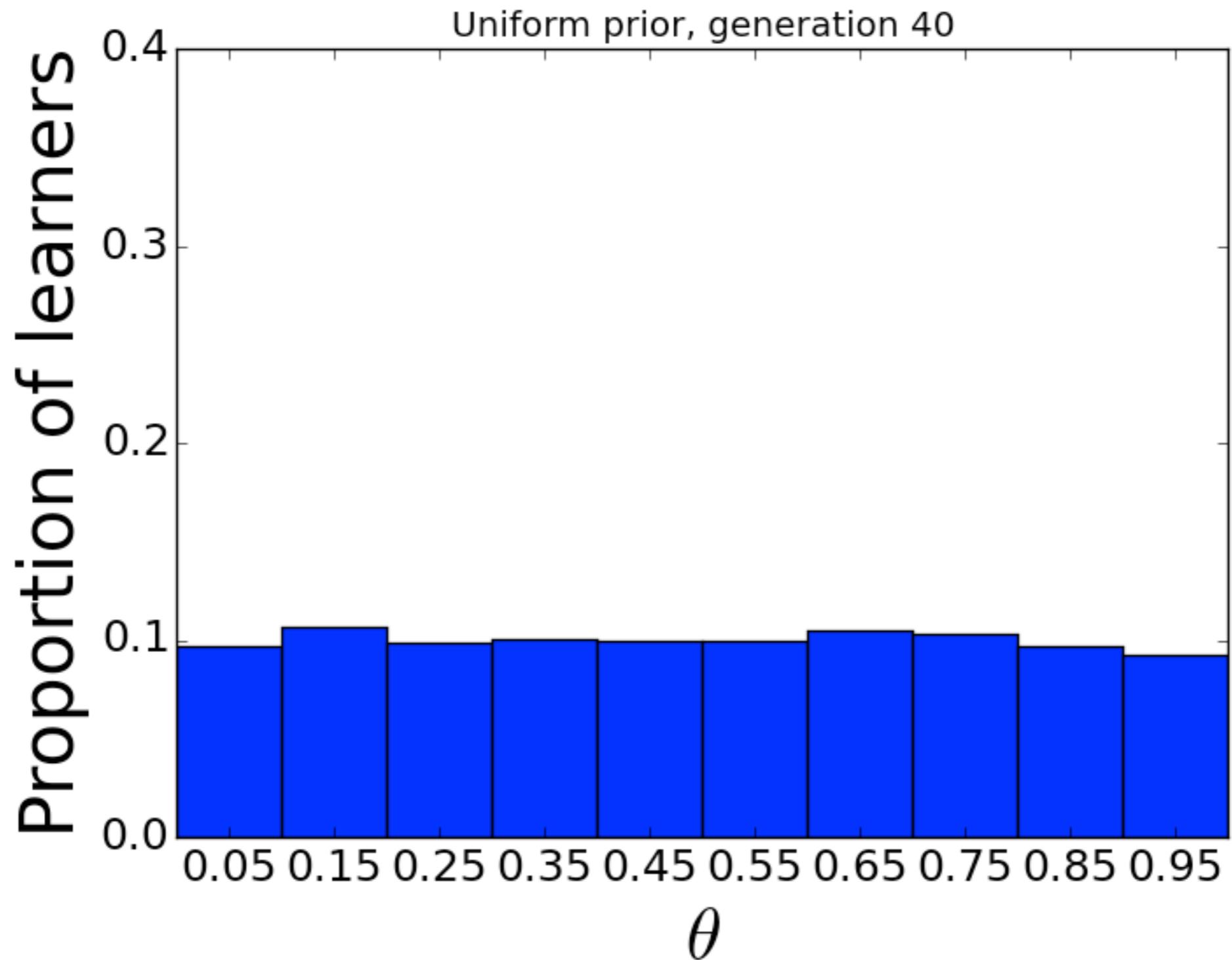
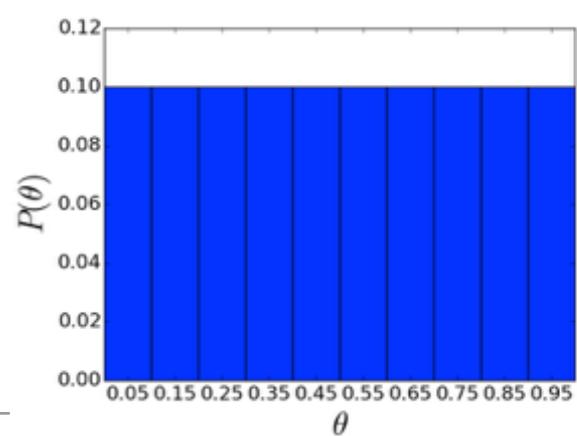
# Watching the prior reveal itself



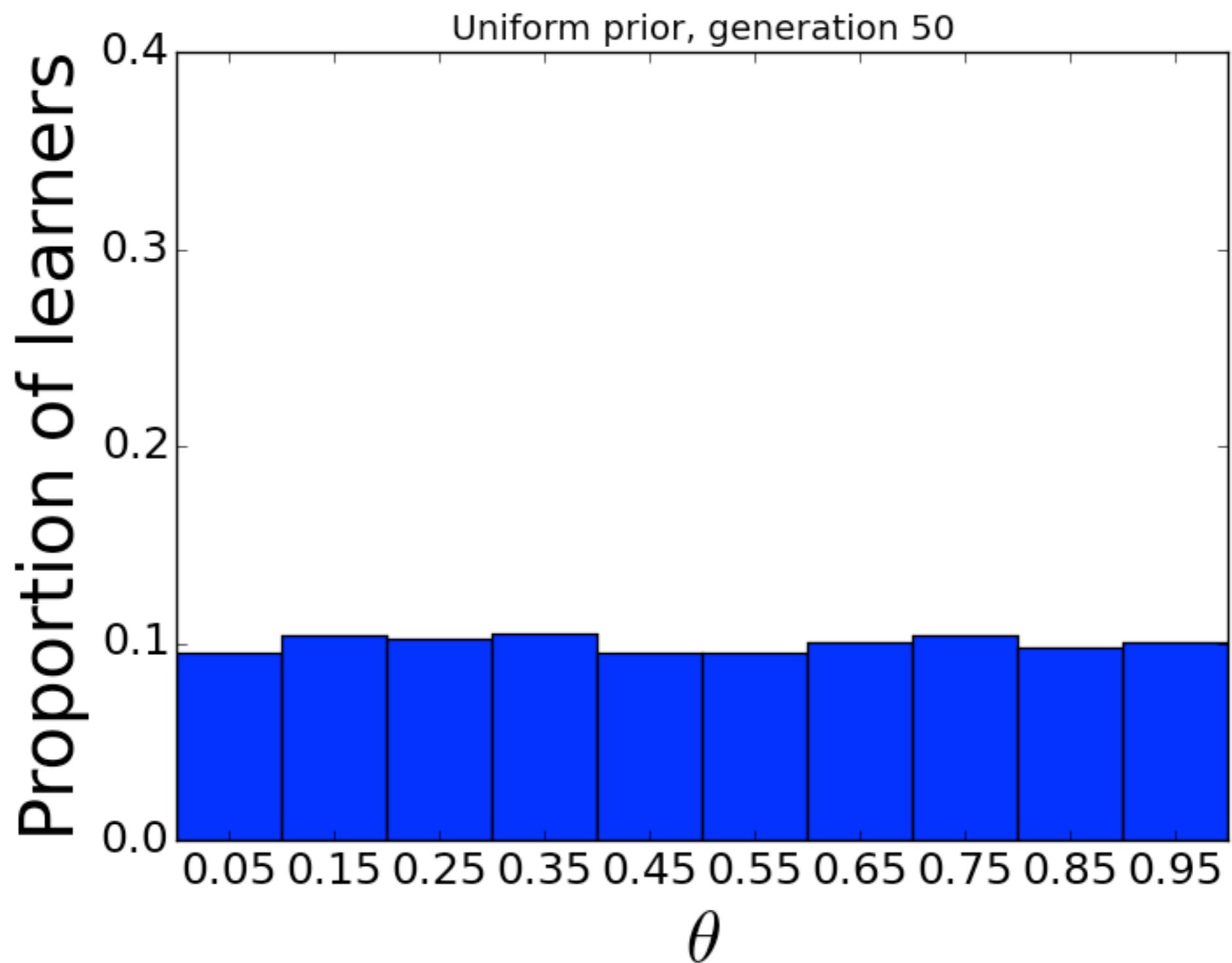
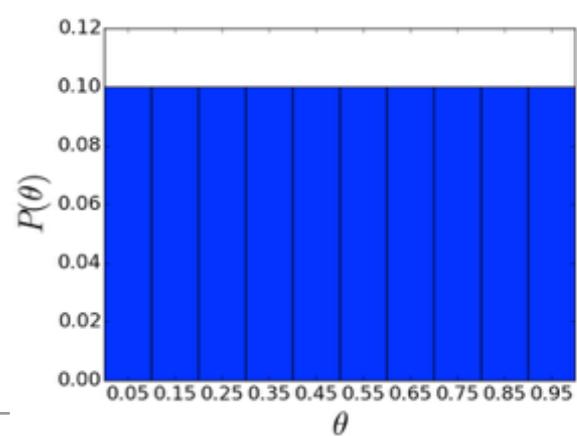
# Watching the prior reveal itself



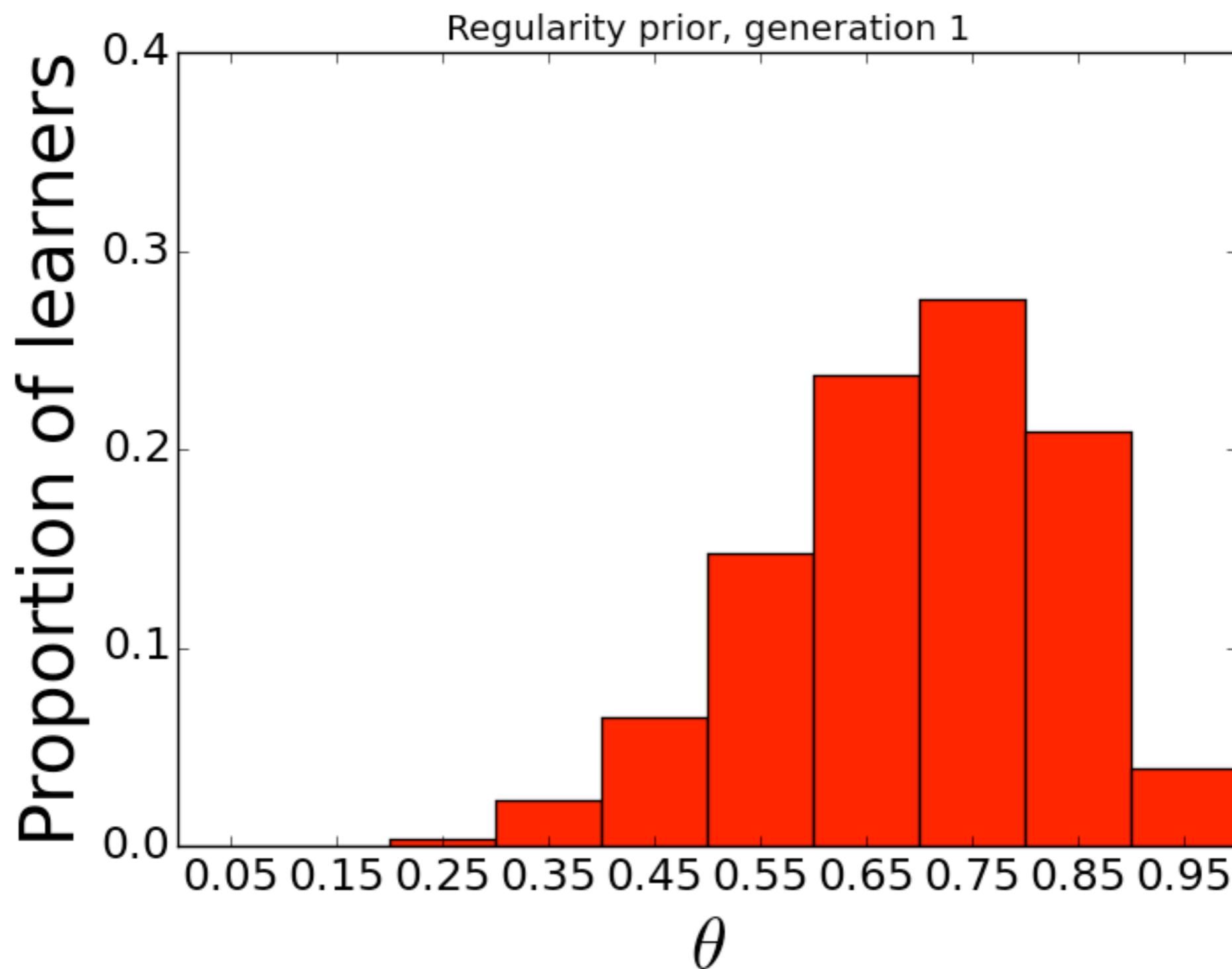
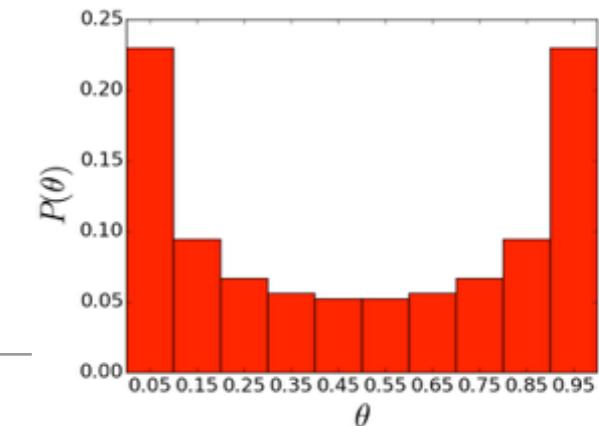
# Watching the prior reveal itself



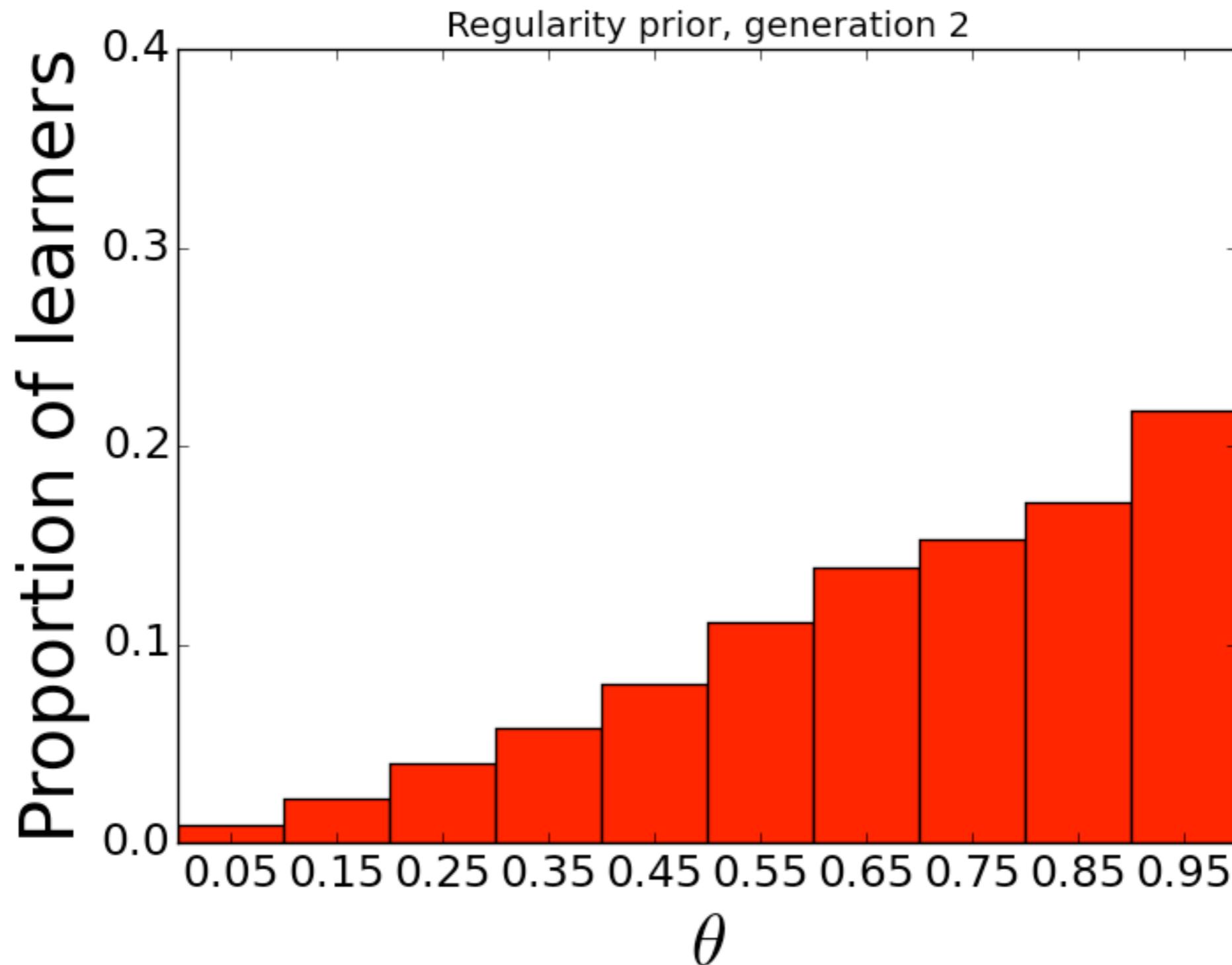
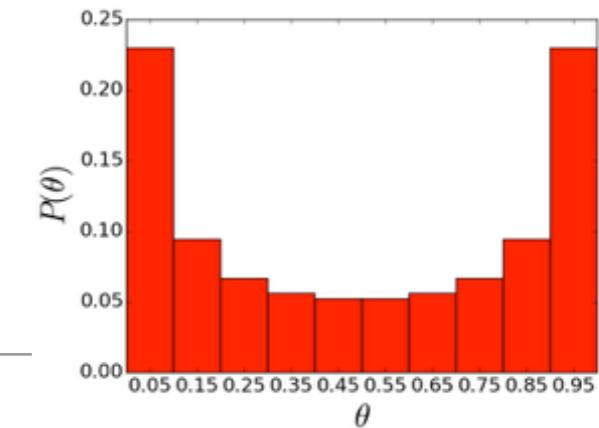
# Watching the prior reveal itself



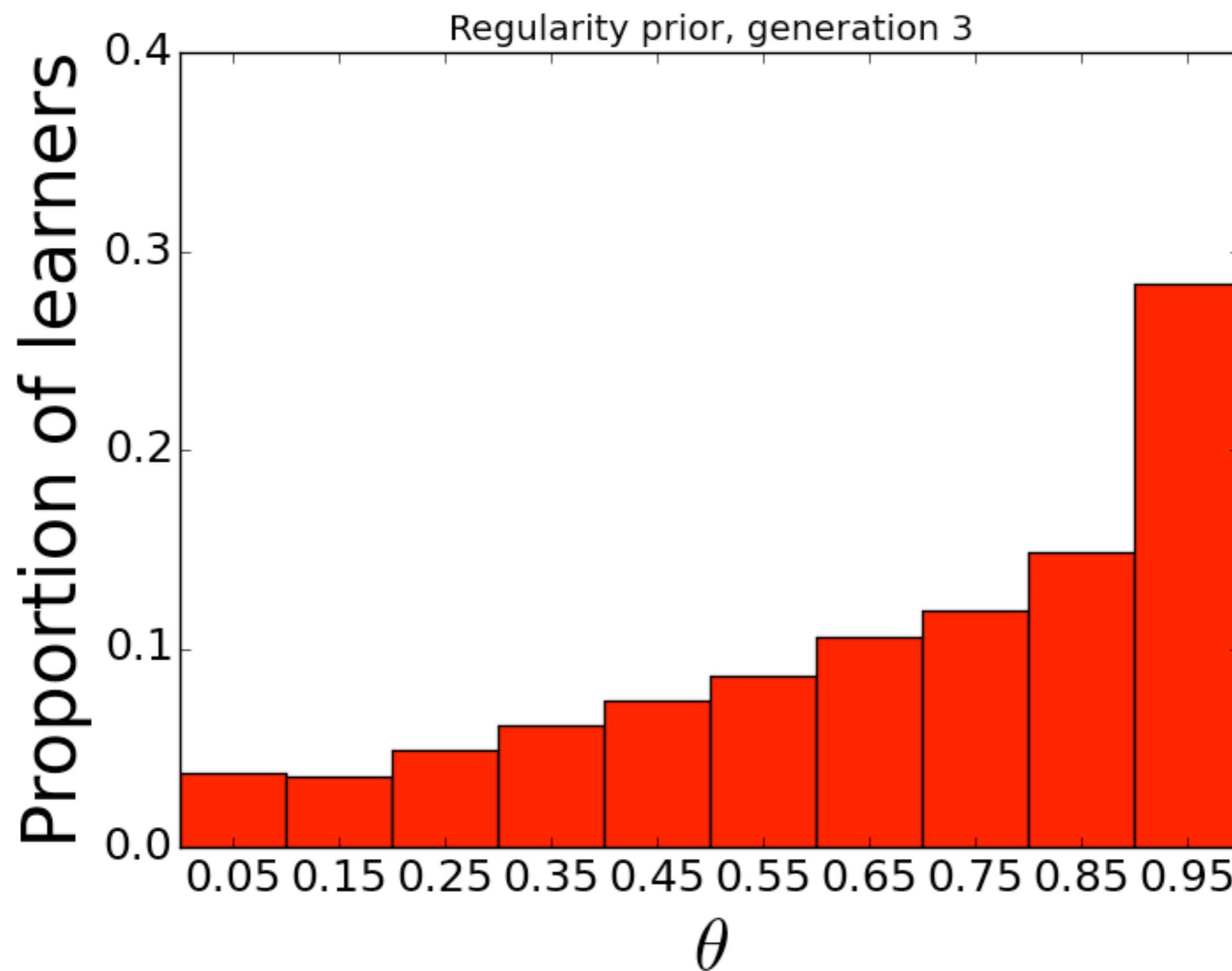
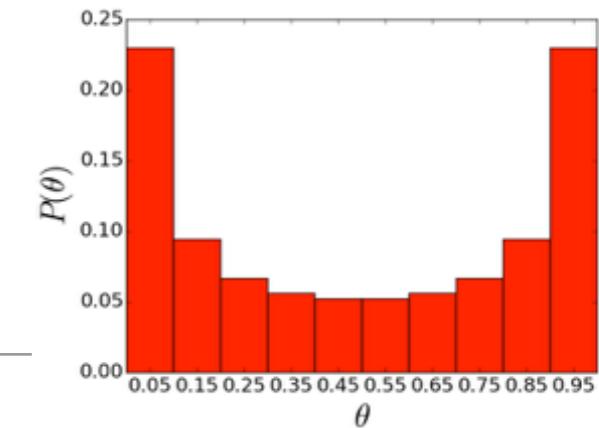
# Watching the prior reveal itself



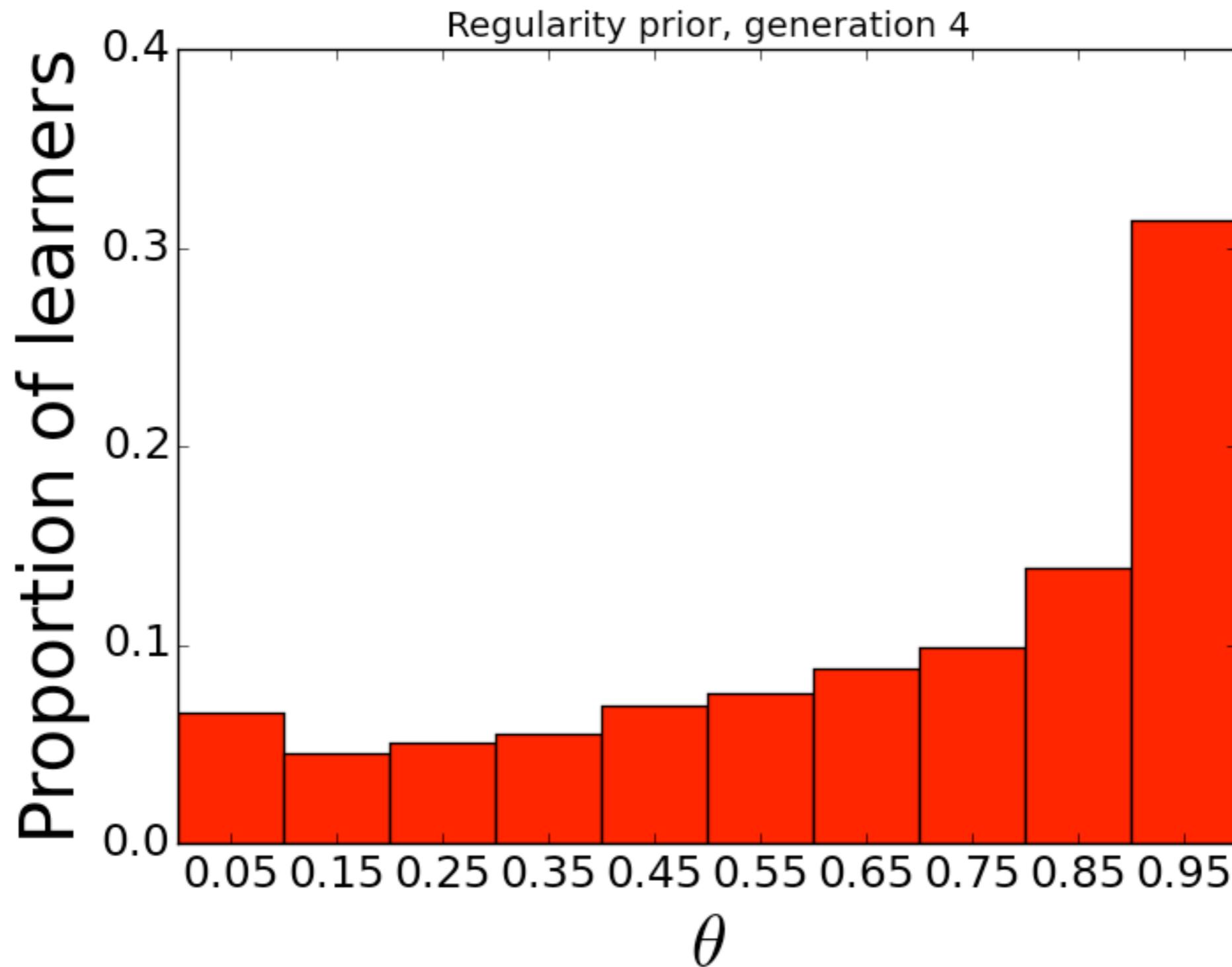
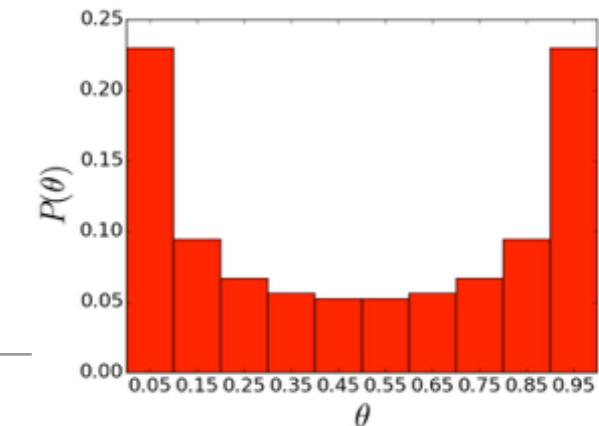
# Watching the prior reveal itself



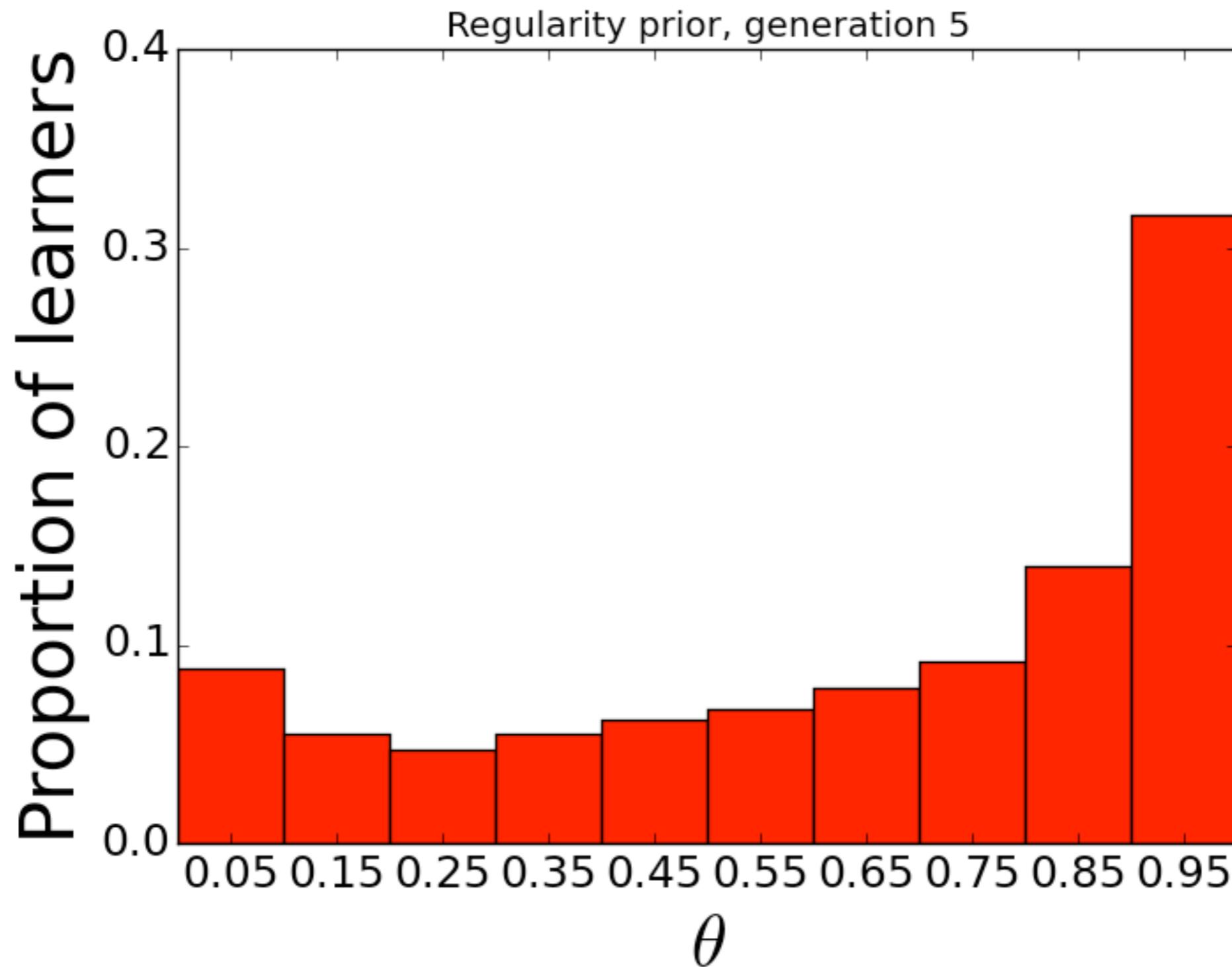
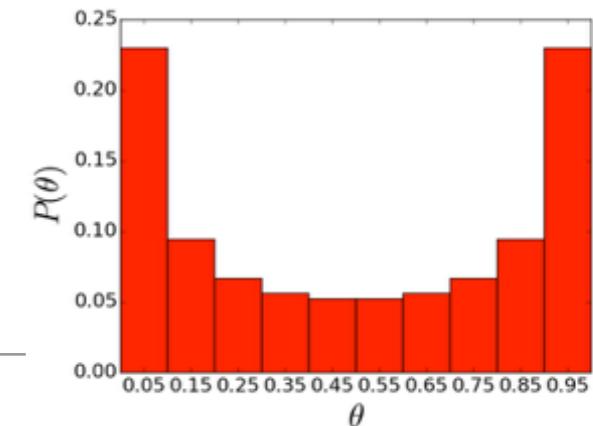
# Watching the prior reveal itself



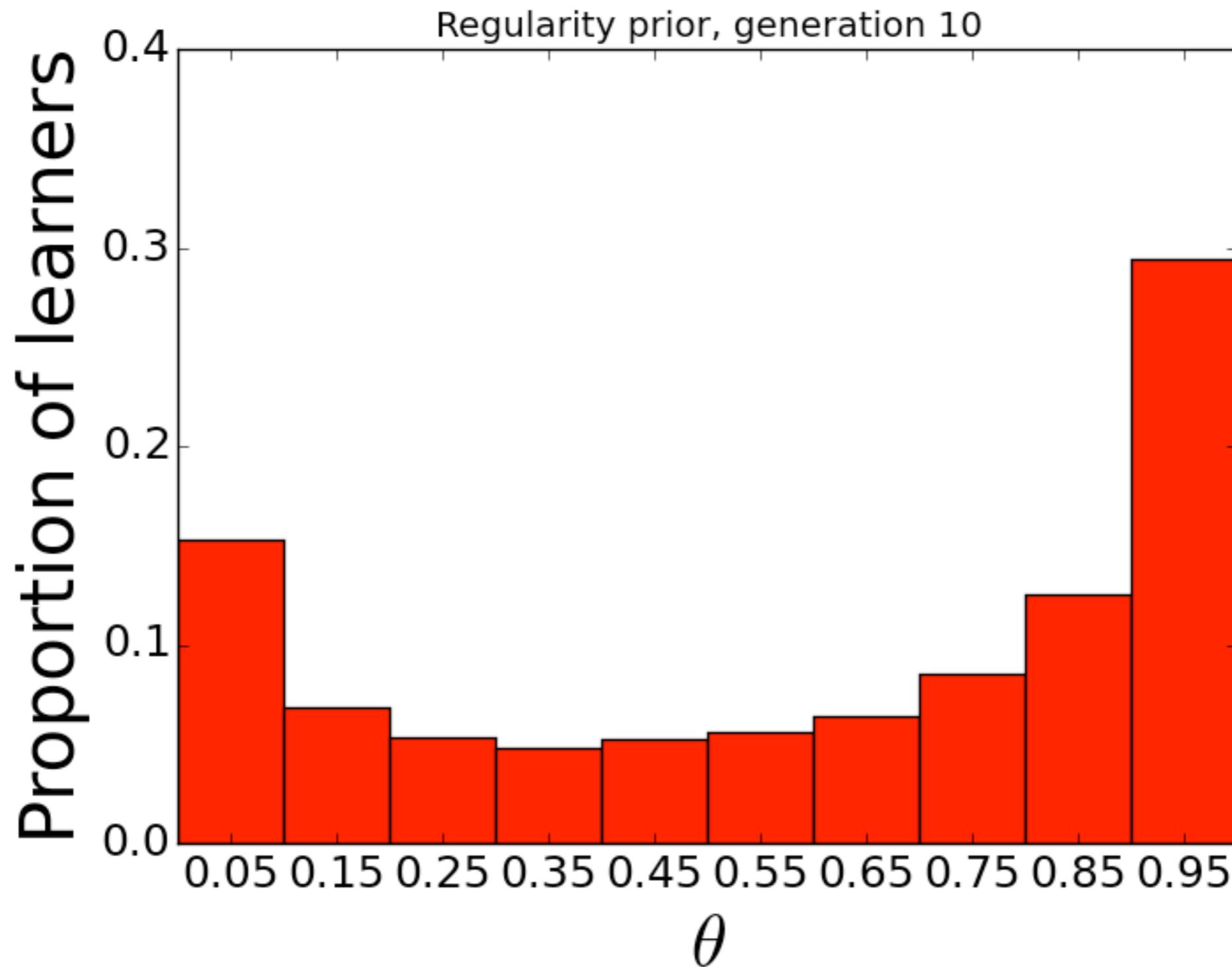
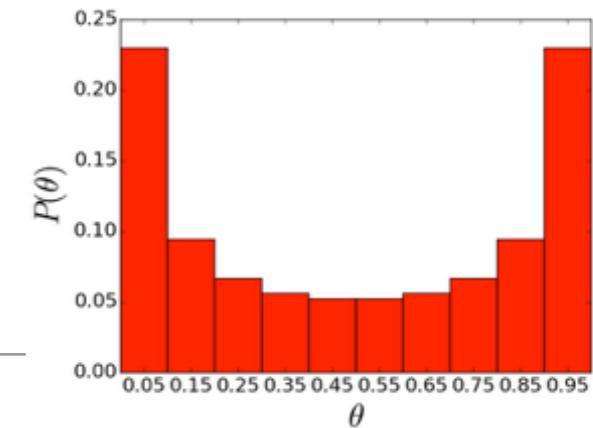
# Watching the prior reveal itself



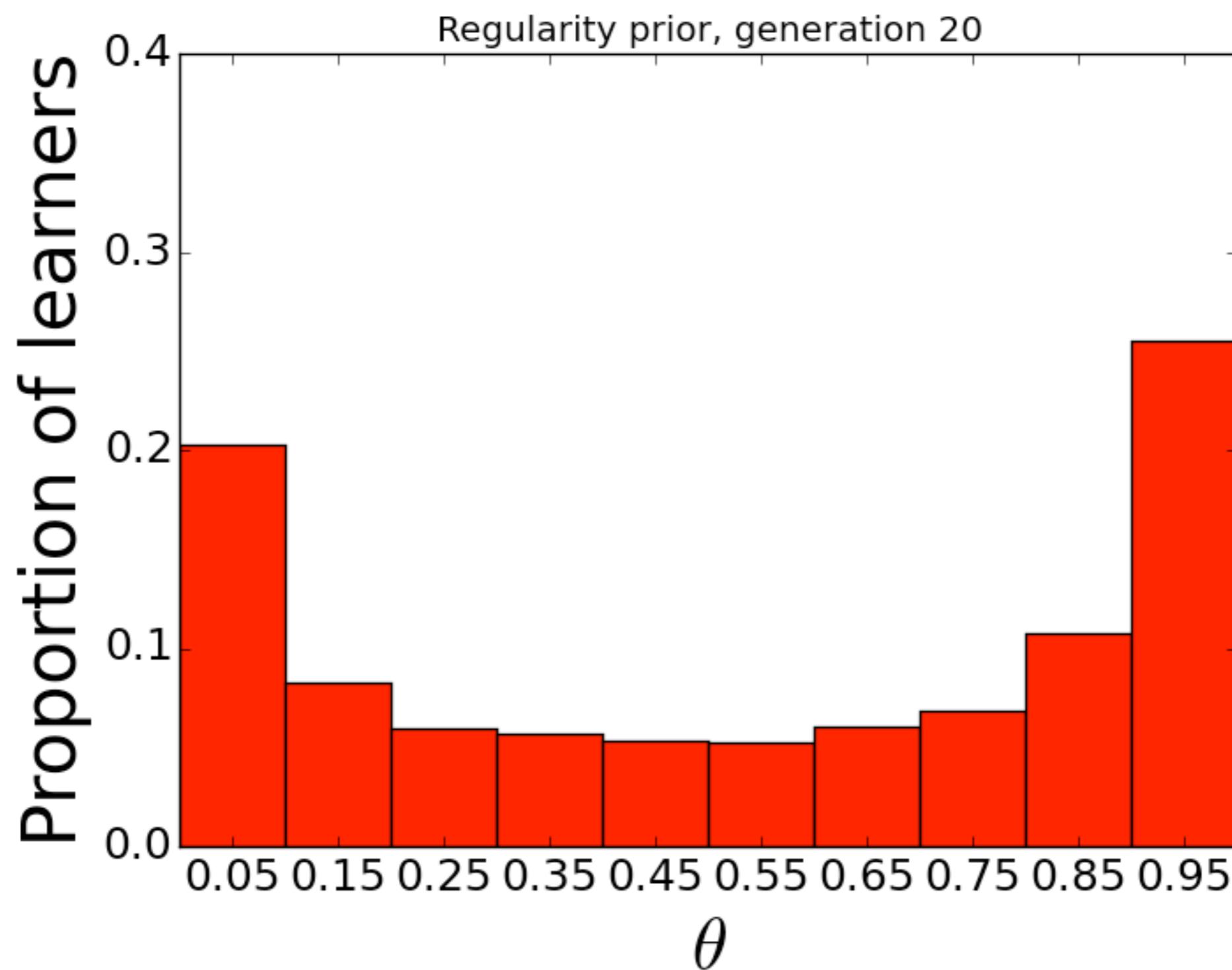
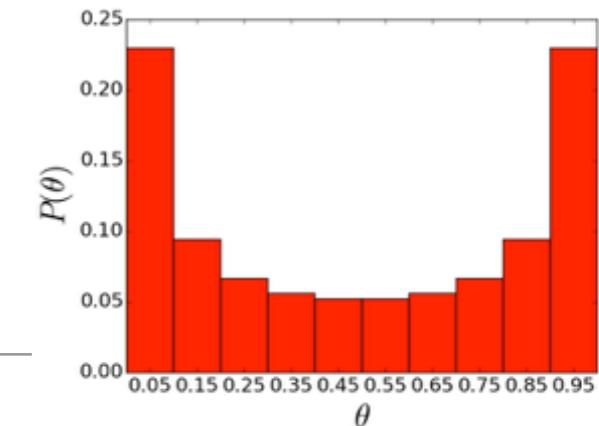
# Watching the prior reveal itself



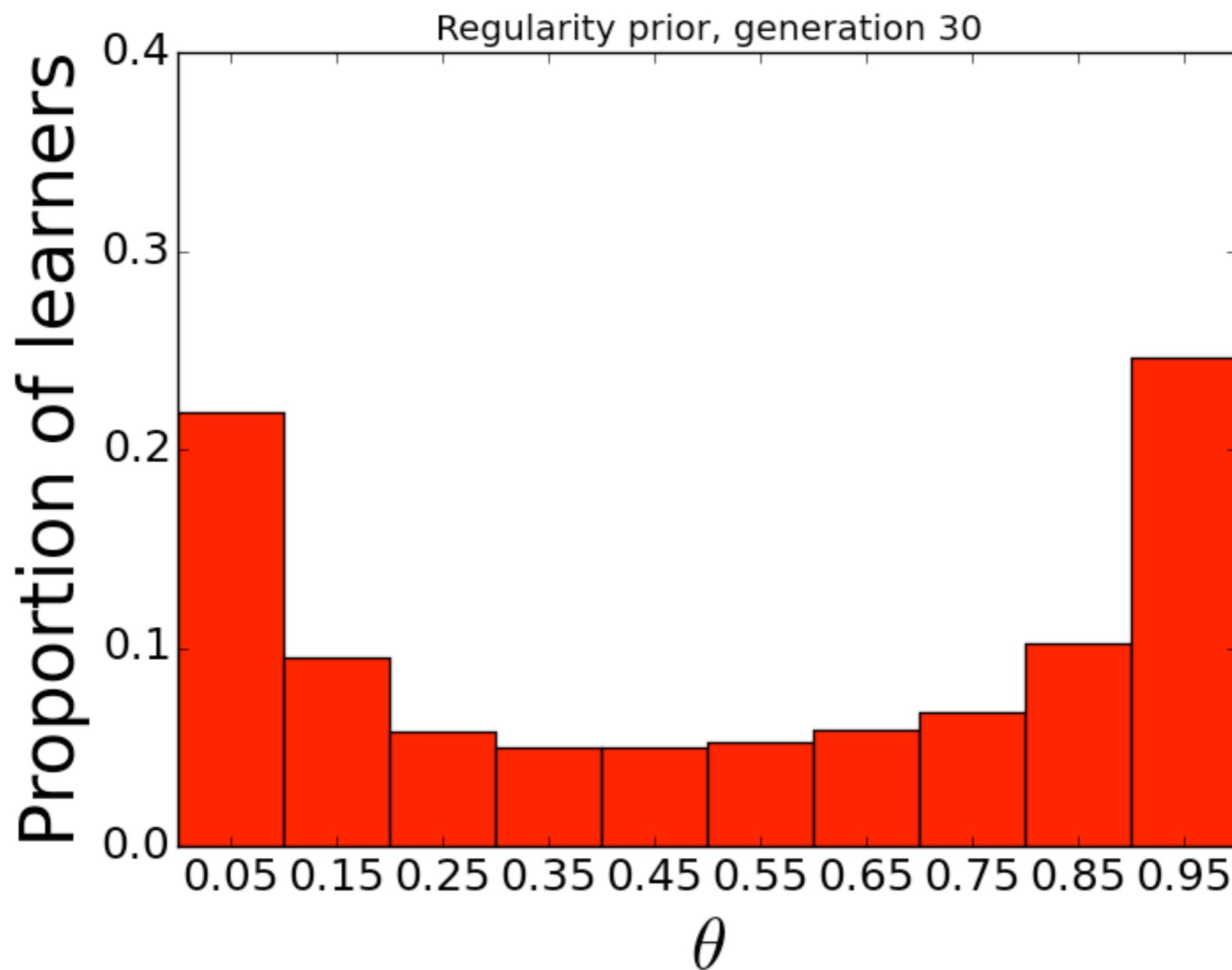
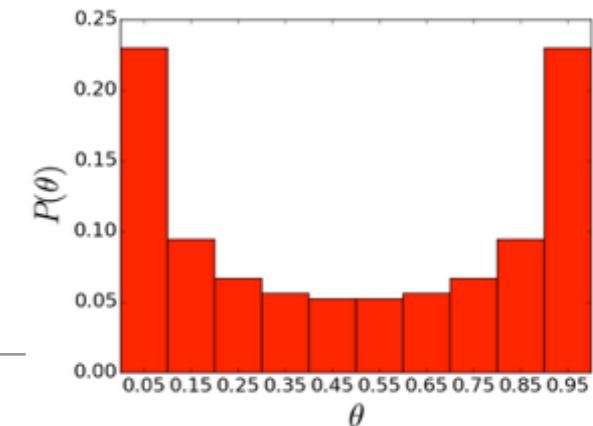
# Watching the prior reveal itself



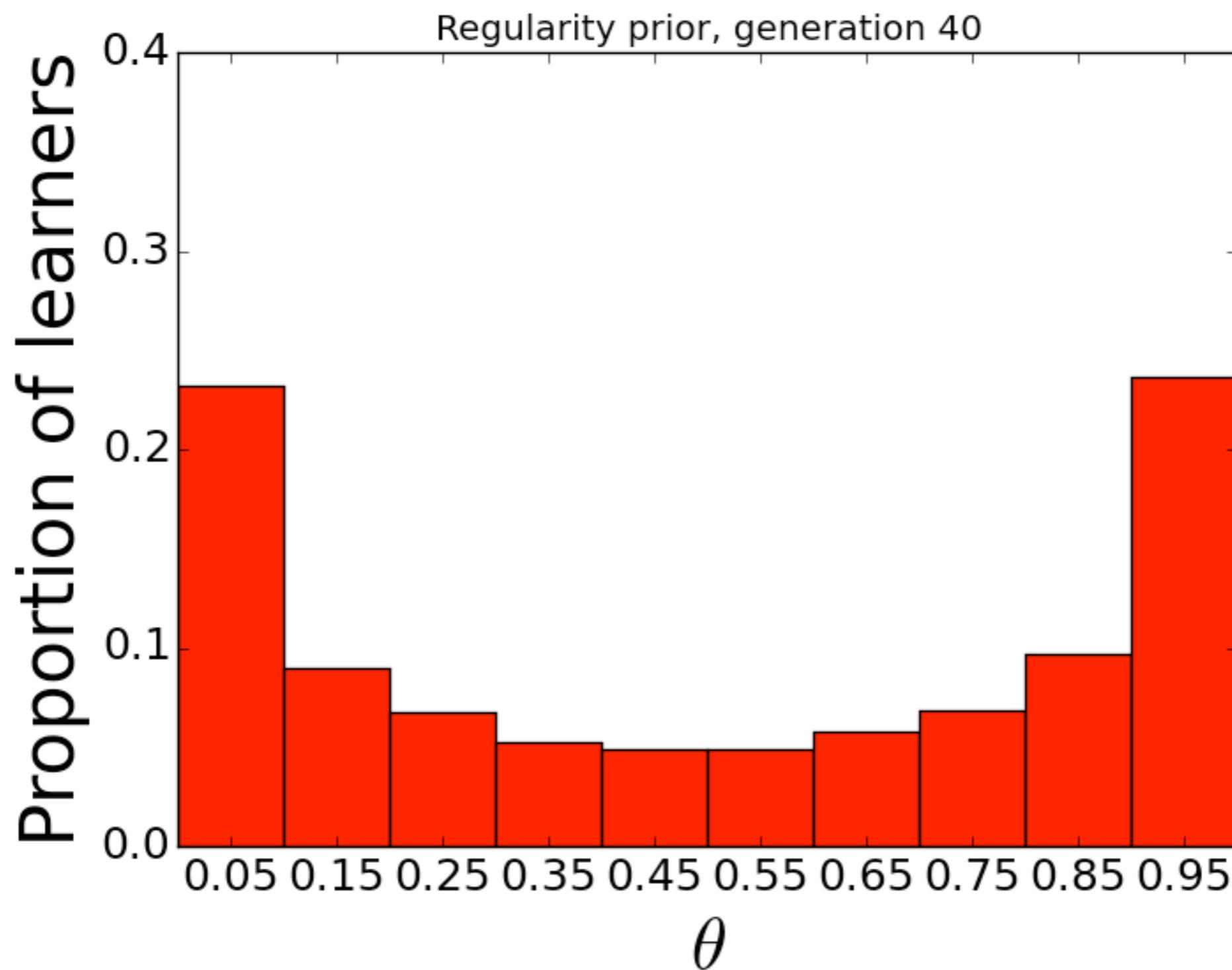
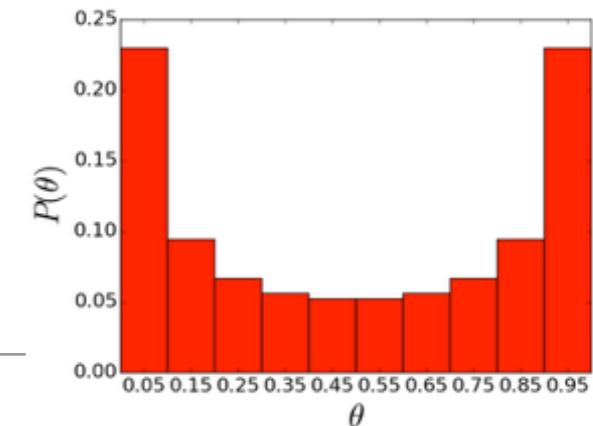
# Watching the prior reveal itself



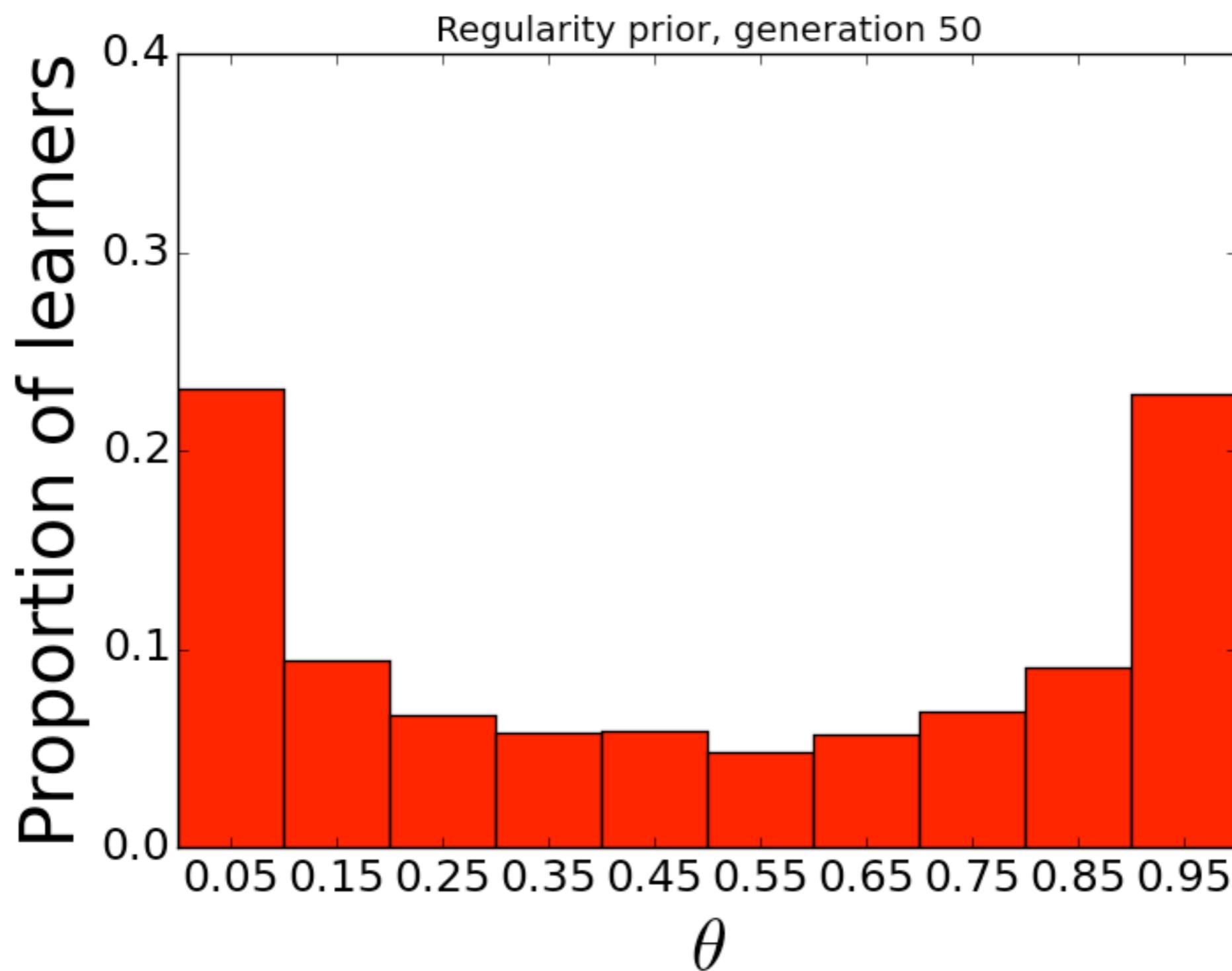
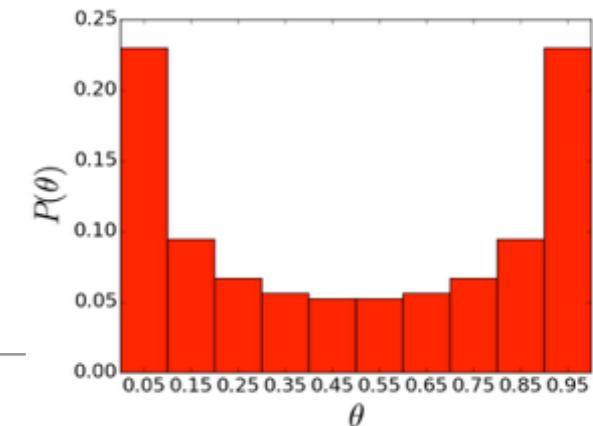
# Watching the prior reveal itself



# Watching the prior reveal itself

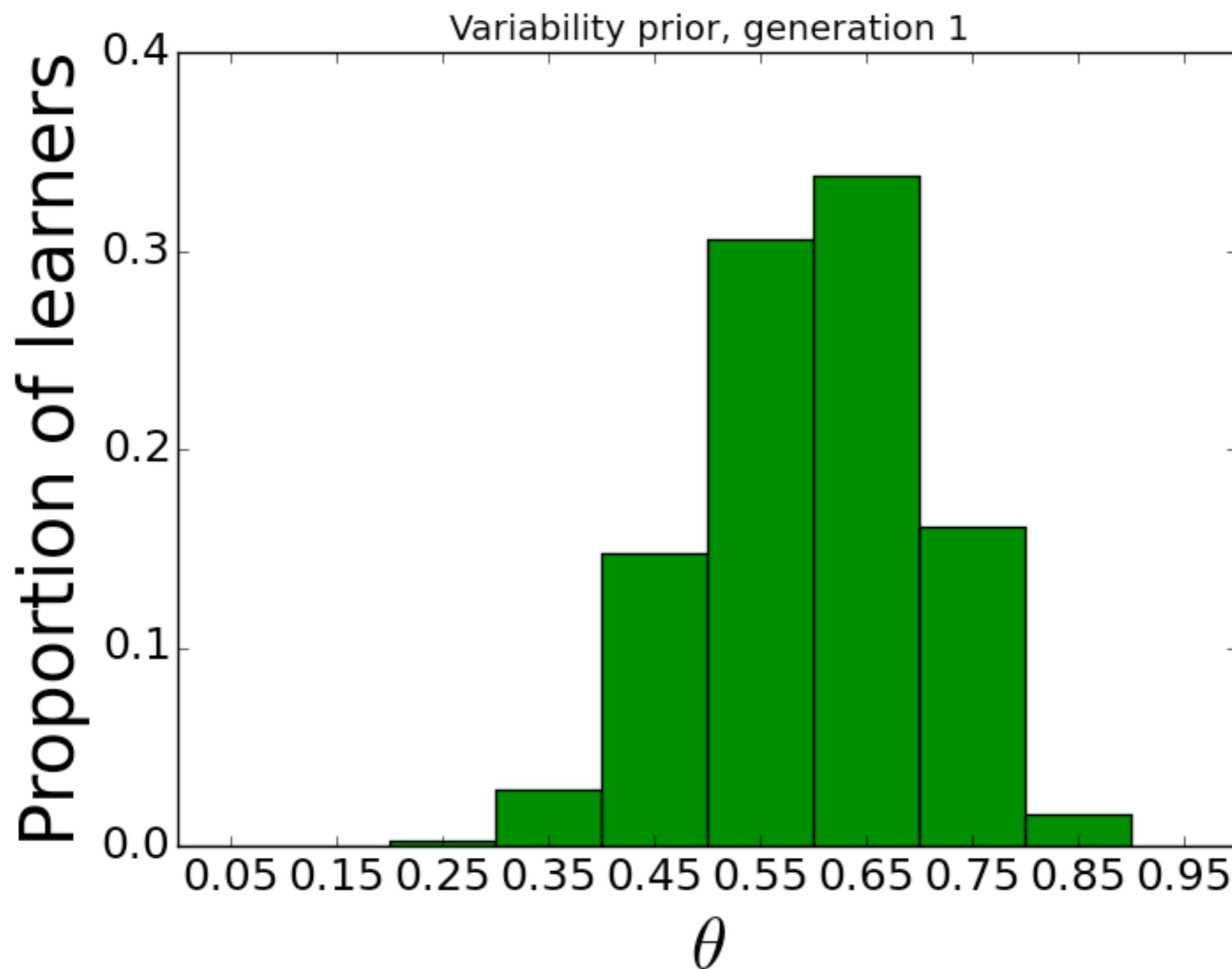
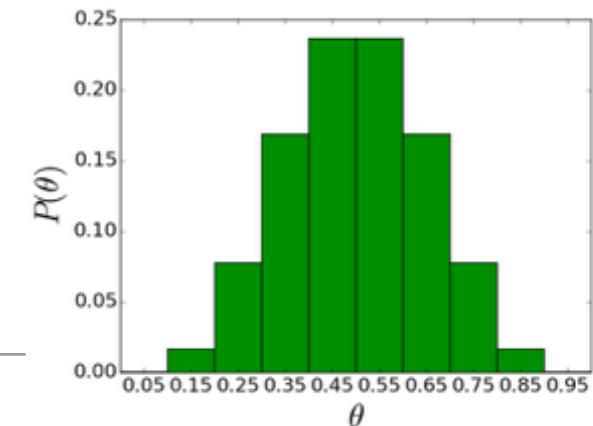


# Watching the prior reveal itself



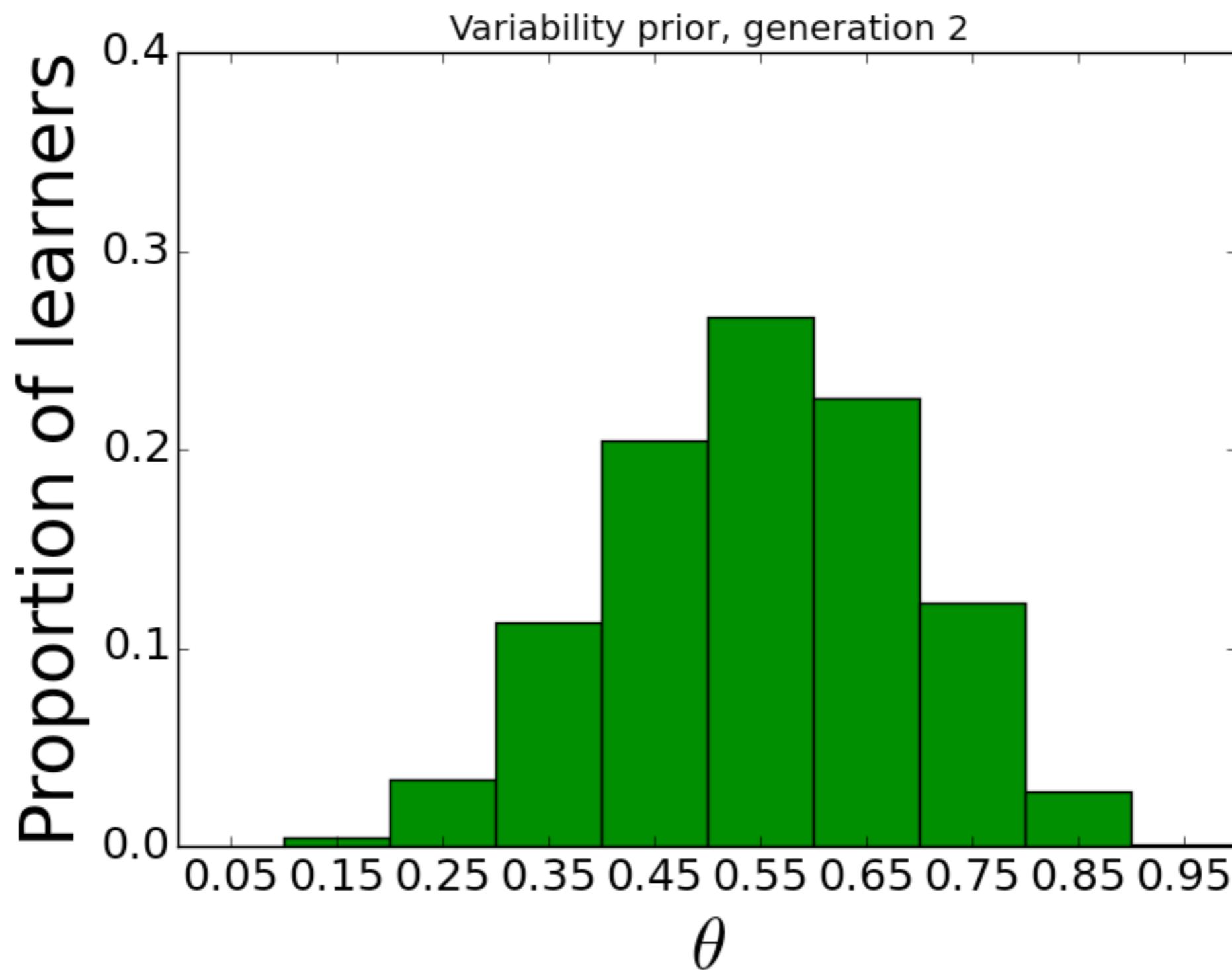
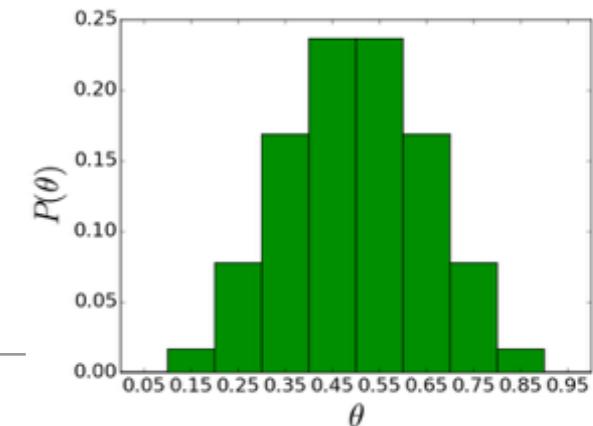
# Watching the prior reveal itself

---



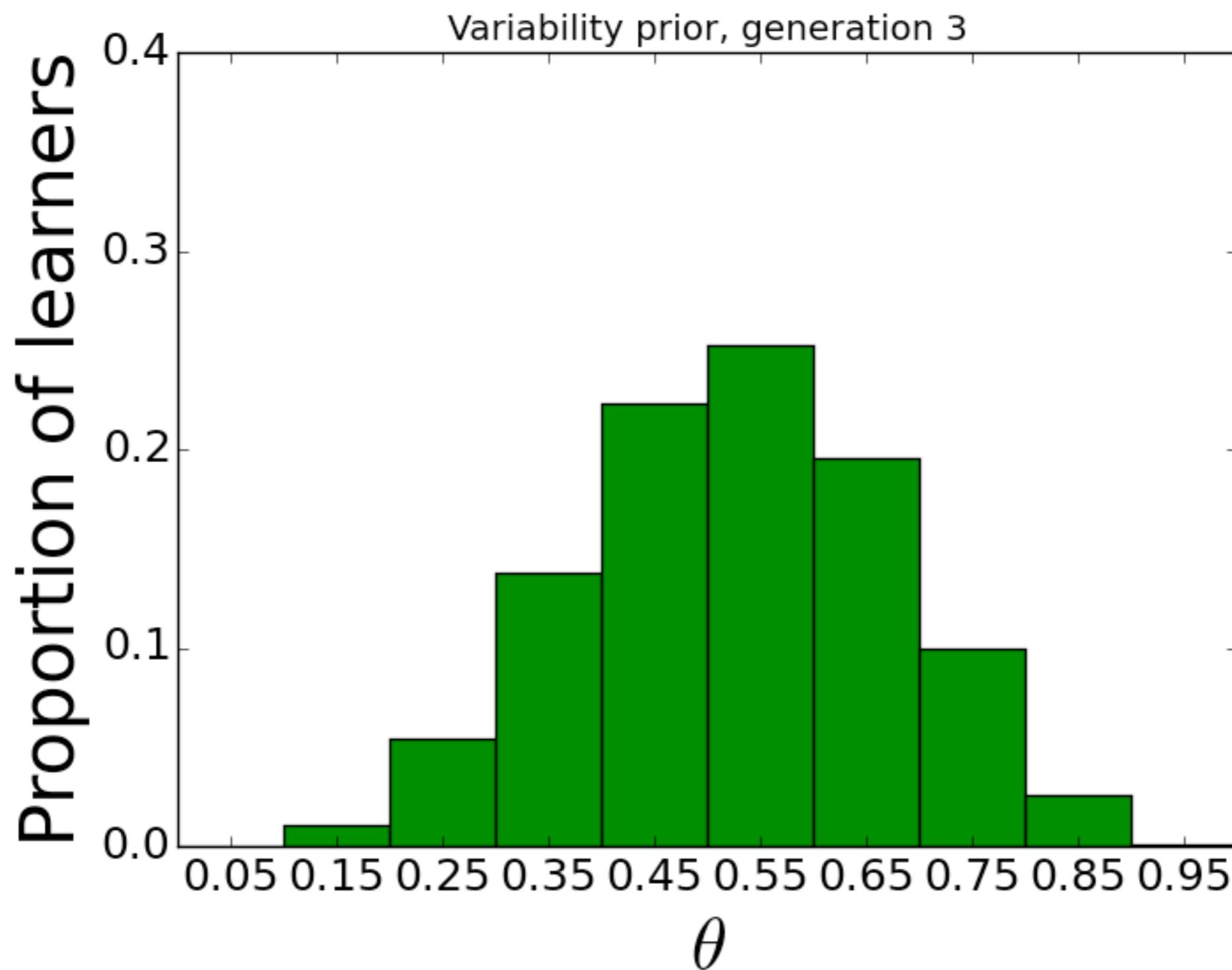
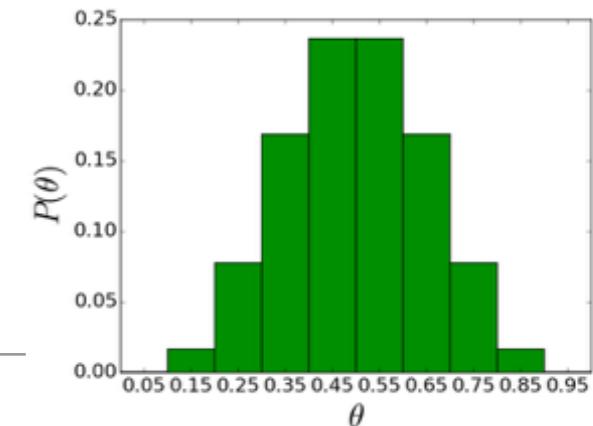
# Watching the prior reveal itself

---



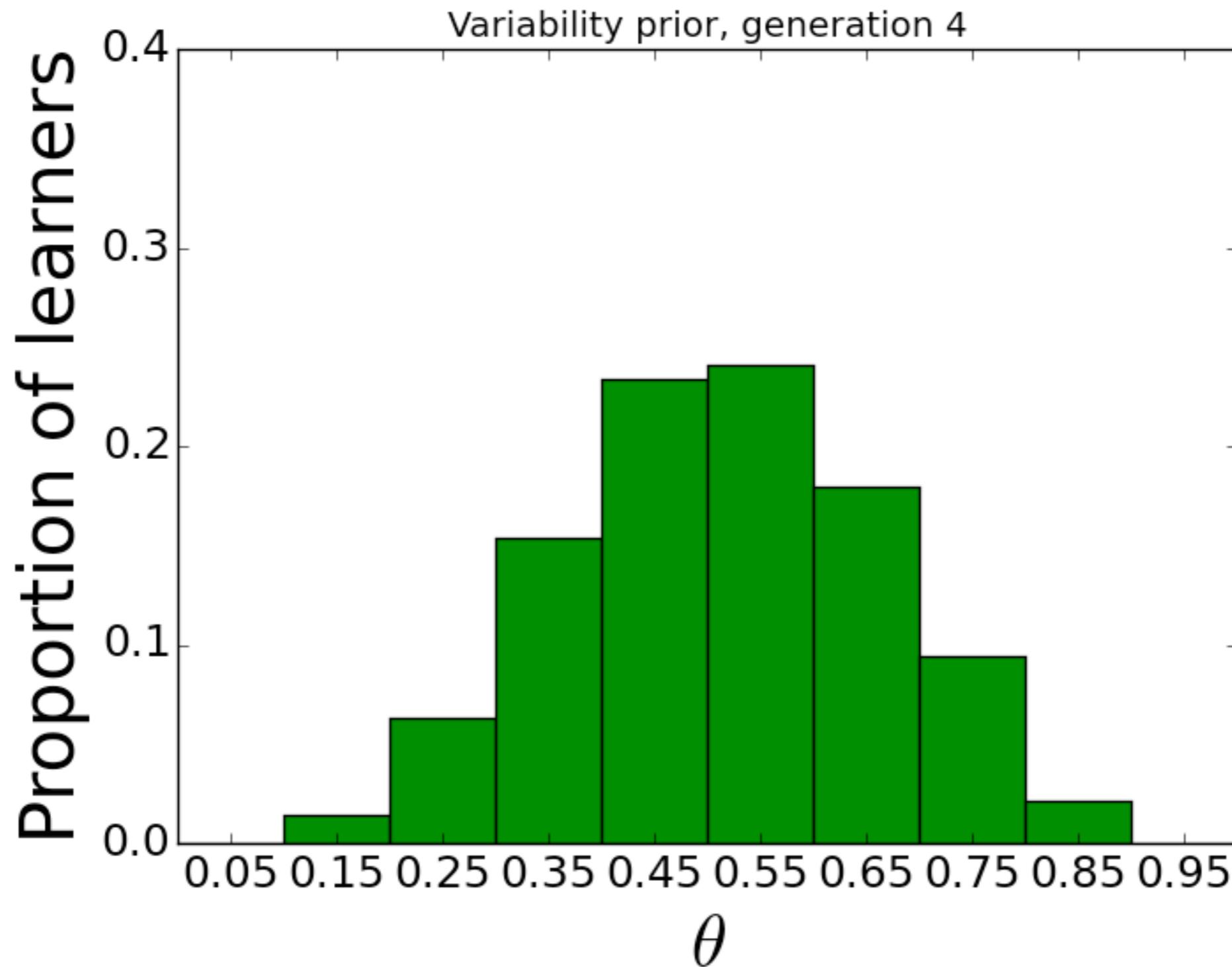
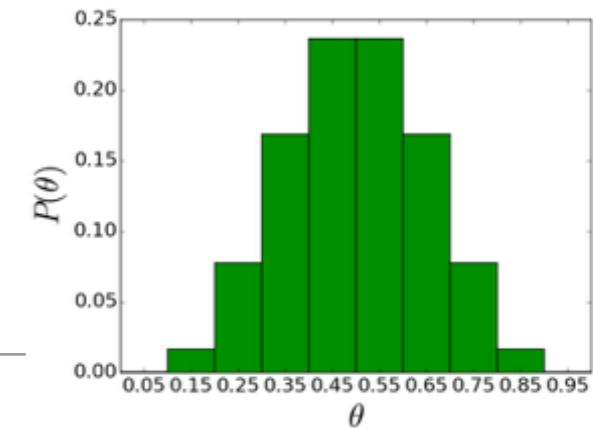
# Watching the prior reveal itself

---



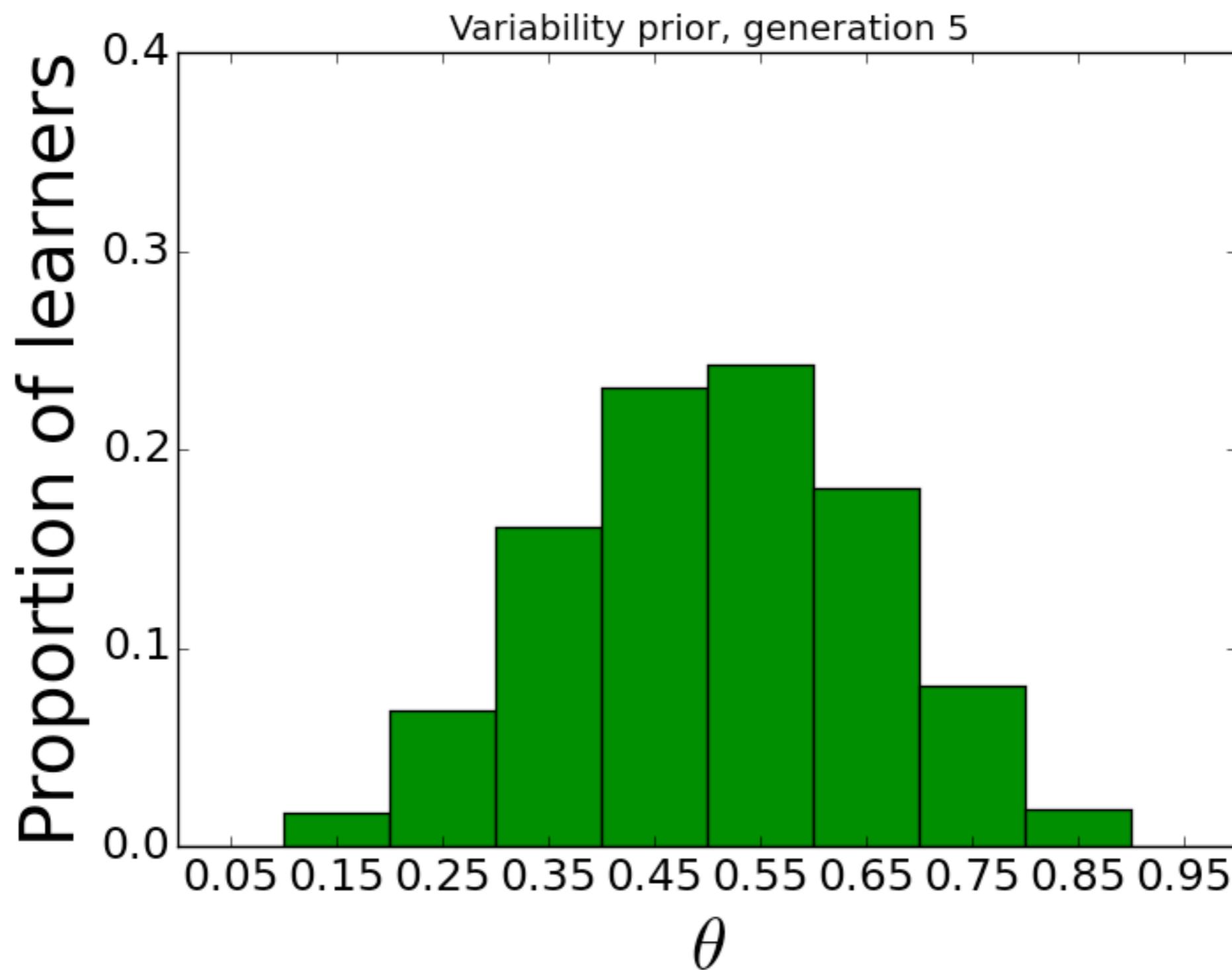
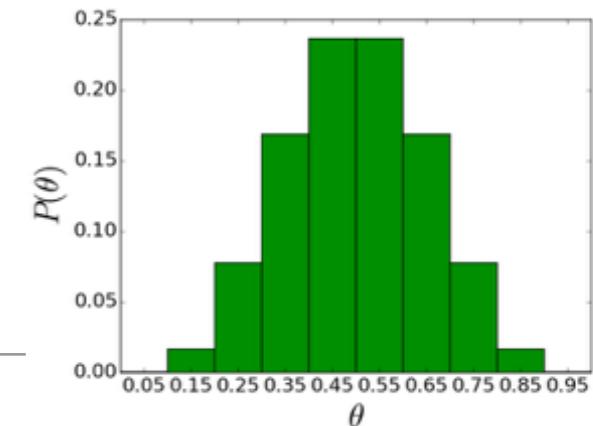
# Watching the prior reveal itself

---

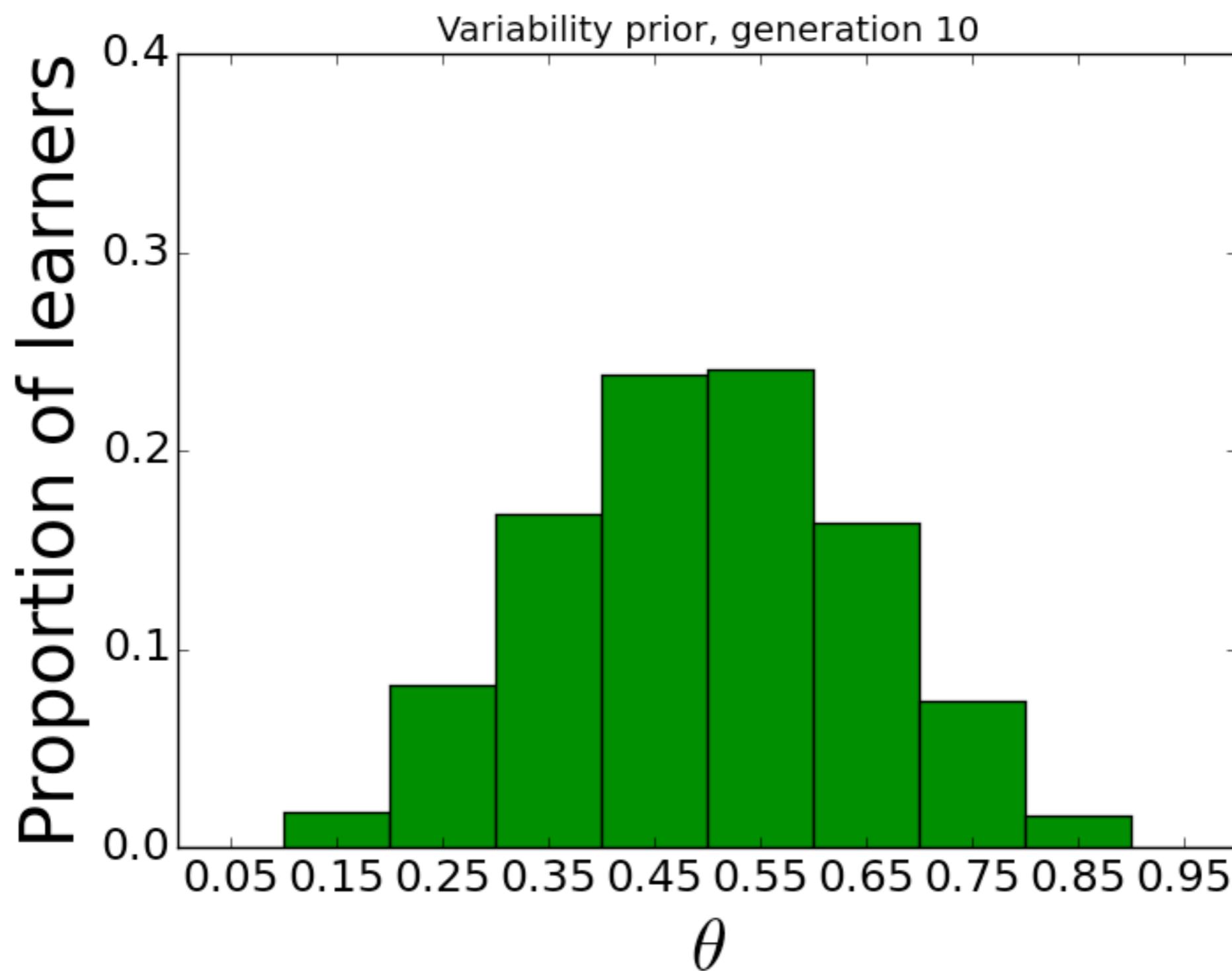
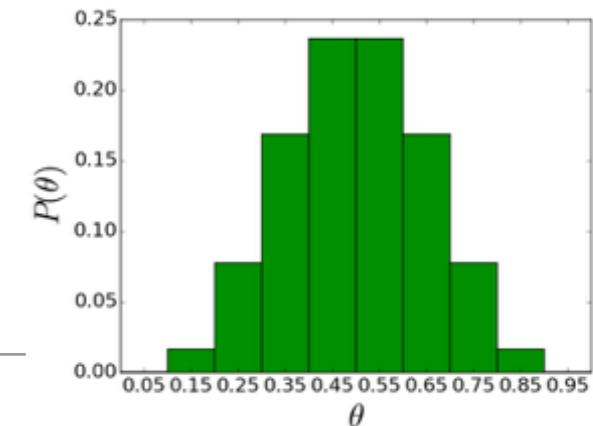


# Watching the prior reveal itself

---



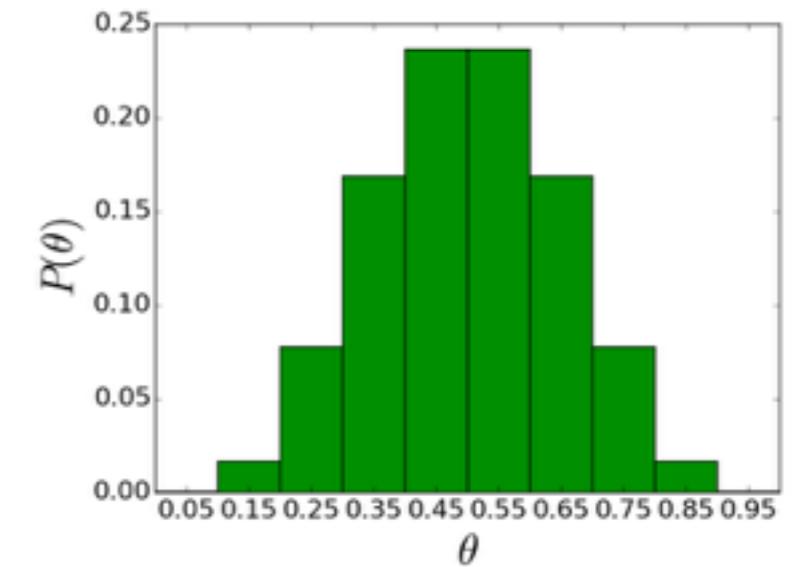
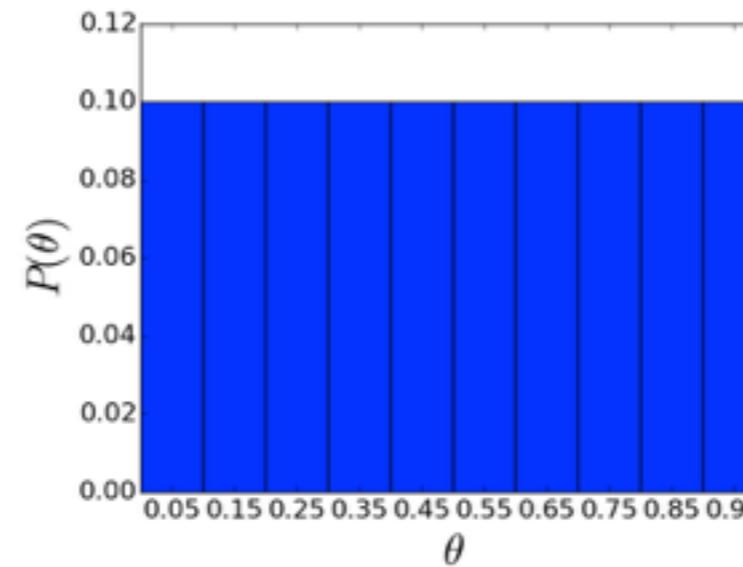
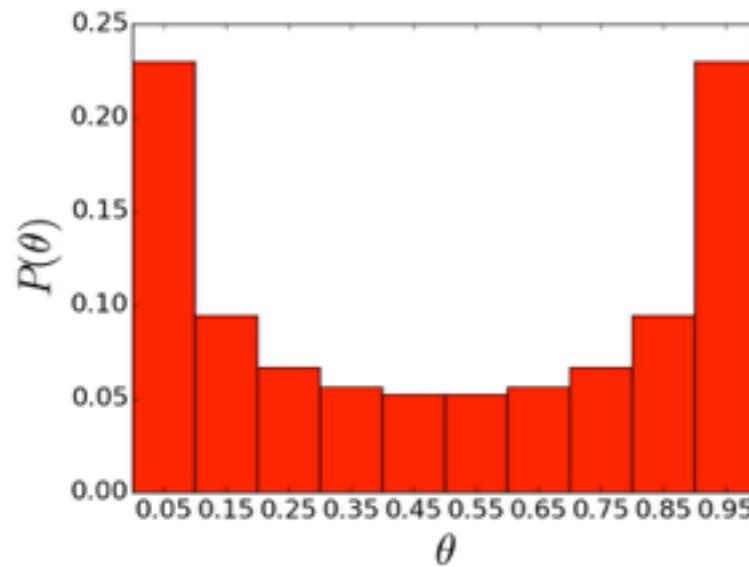
# Watching the prior reveal itself



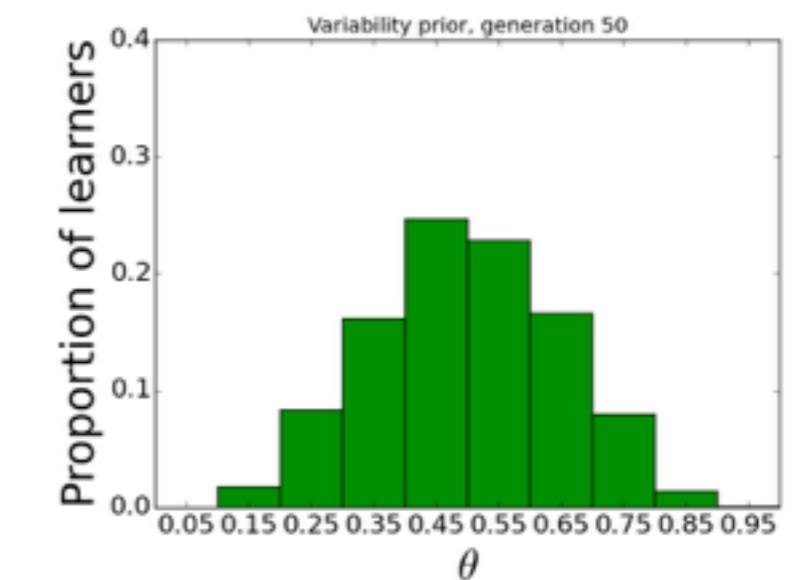
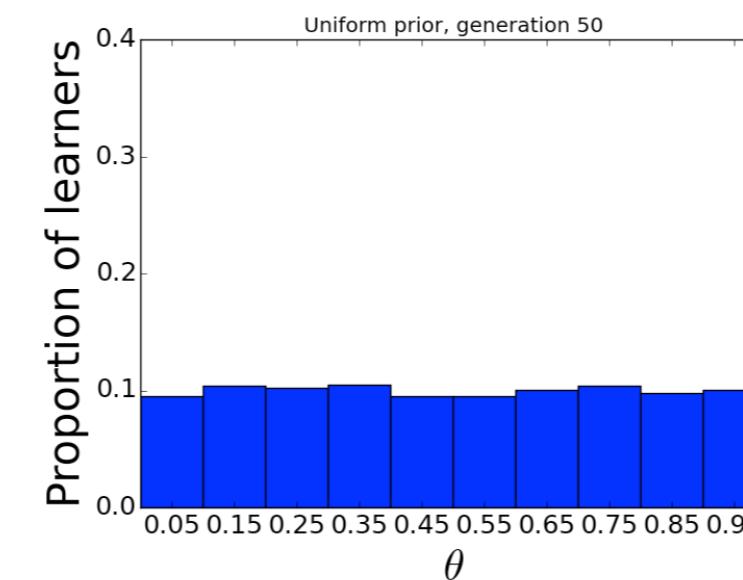
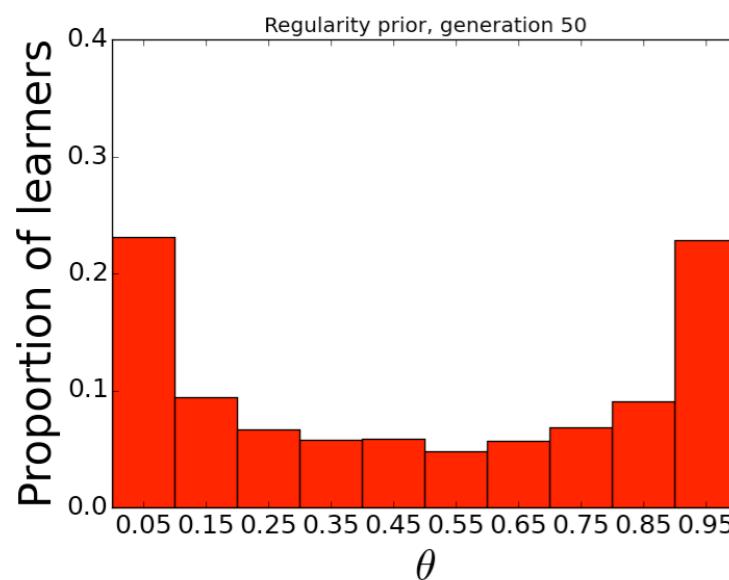
# Culture converges to the prior

---

- Priors



- Distribution of languages after 50 generations



# What is the relationship between languages and language learners here?

---

- The types of languages we see in the world should:
  - A. be completely unconstrained by the biases of language learners
  - B. reflect the biases of language learners, but in an interestingly complex way (weak biases have strong effects, bottleneck matters, etc)
  - C. directly reflect the biases of language learners

# Hang on a minute...

---

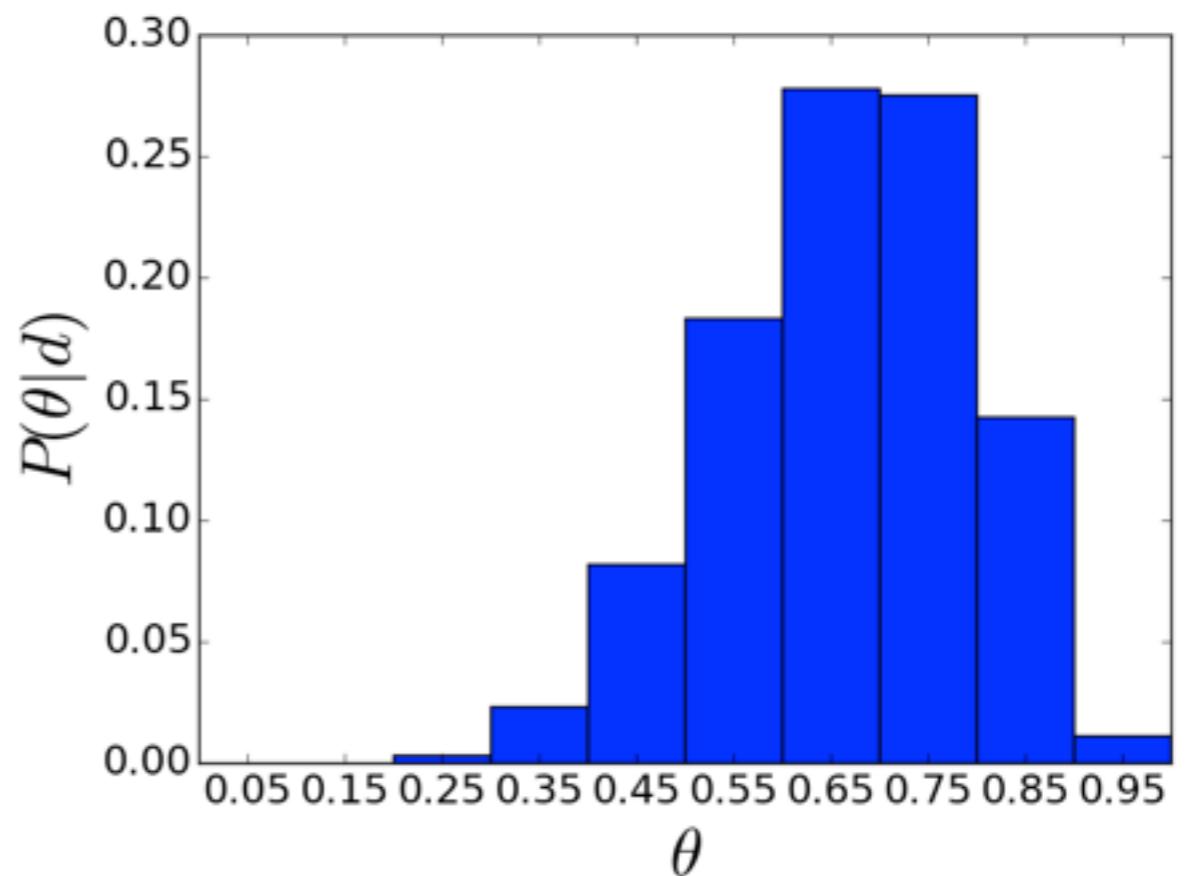
- This runs counter to the results from my simulation in Kirby (2002)
  - I argued that it was the bottleneck that was driving adaptation of the language
- It also runs counter to the spirit of all the stuff I have been saying throughout this course!
  - I argued that cultural evolution has something important to add
  - If prior bias is what is innate to the learner, then the Griffiths & Kalish result suggests that the universal properties of language are just a straightforward reflection of innateness
  - Hmm...

# Some subtleties in the model

---

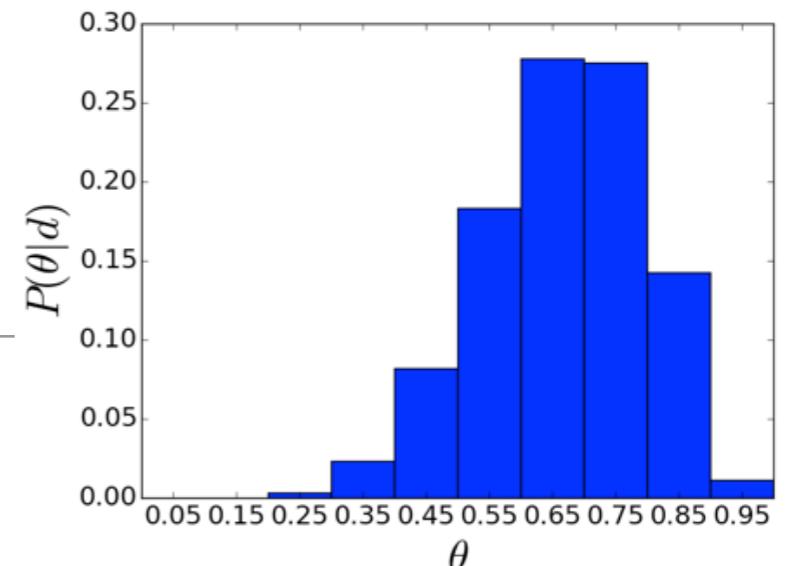
- Kirby, Dowman & Griffiths (2007): tried to square the Bayesian model with what we **thought** we knew about cultural evolution of language
- Whole thing revolves around a very subtle point
  - How do you decide, given the posterior, which language to select?

$$P(h|d) = \frac{P(d|h)P(h)}{P(d)}$$



# Sampling vs. MAP

---



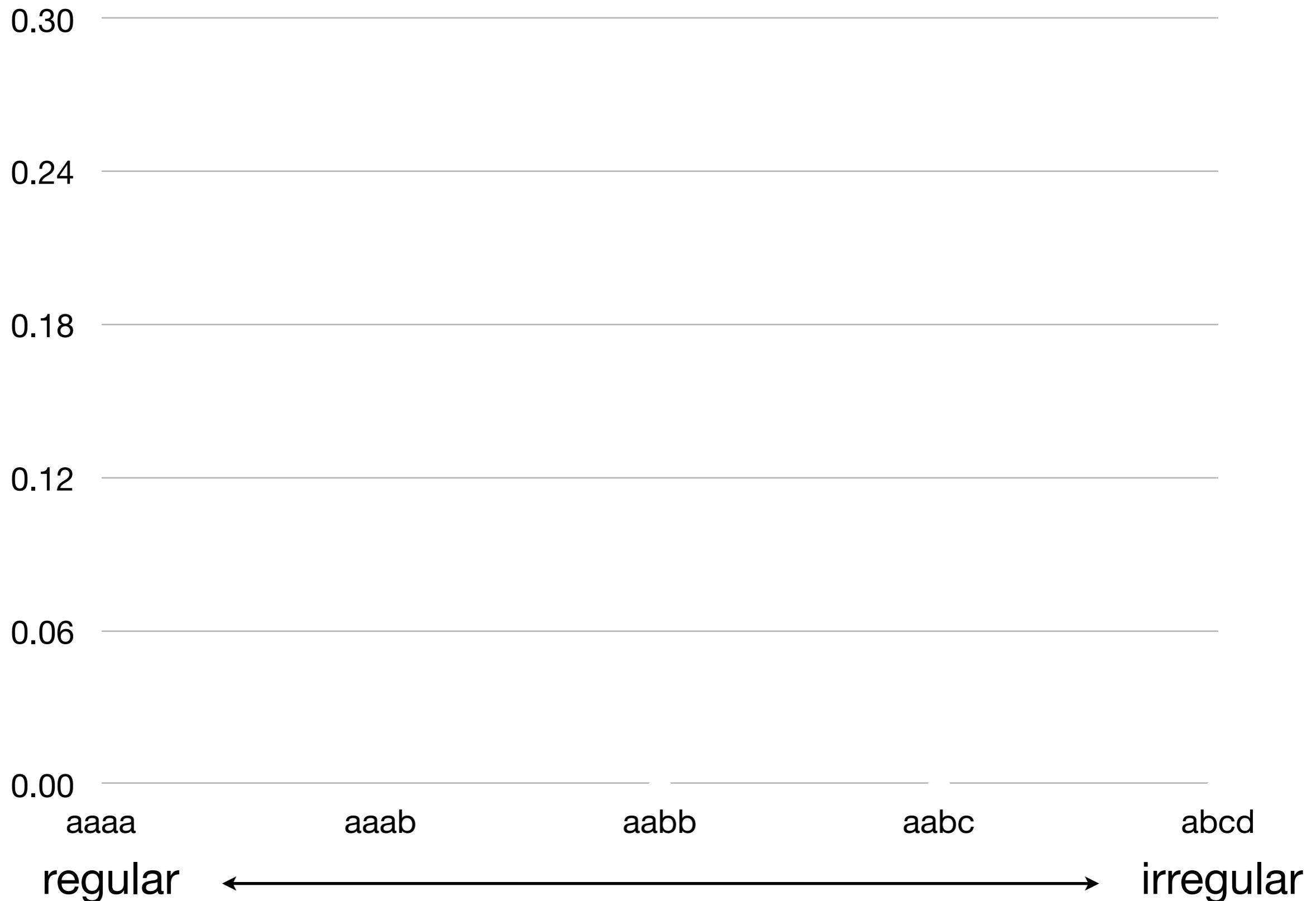
- There are (at least) two sensible choices:
  - Sampling: given a particular distribution of probabilities, pick your hypothesis from the distribution proportionally.  
(If it's ten times more likely to be language A than language B, 10% of the time pick language B)
  - MAP: given a particular distribution of probabilities, pick the best. This is called the maximum a-posteriori (MAP) hypothesis  
(If it's more likely to be language A than language B, pick language A)
- Griffith & Kalish (2007) were using *sampling*. Kirby et al. (2007) tried MAP.

# A final model: the evolution of regularity

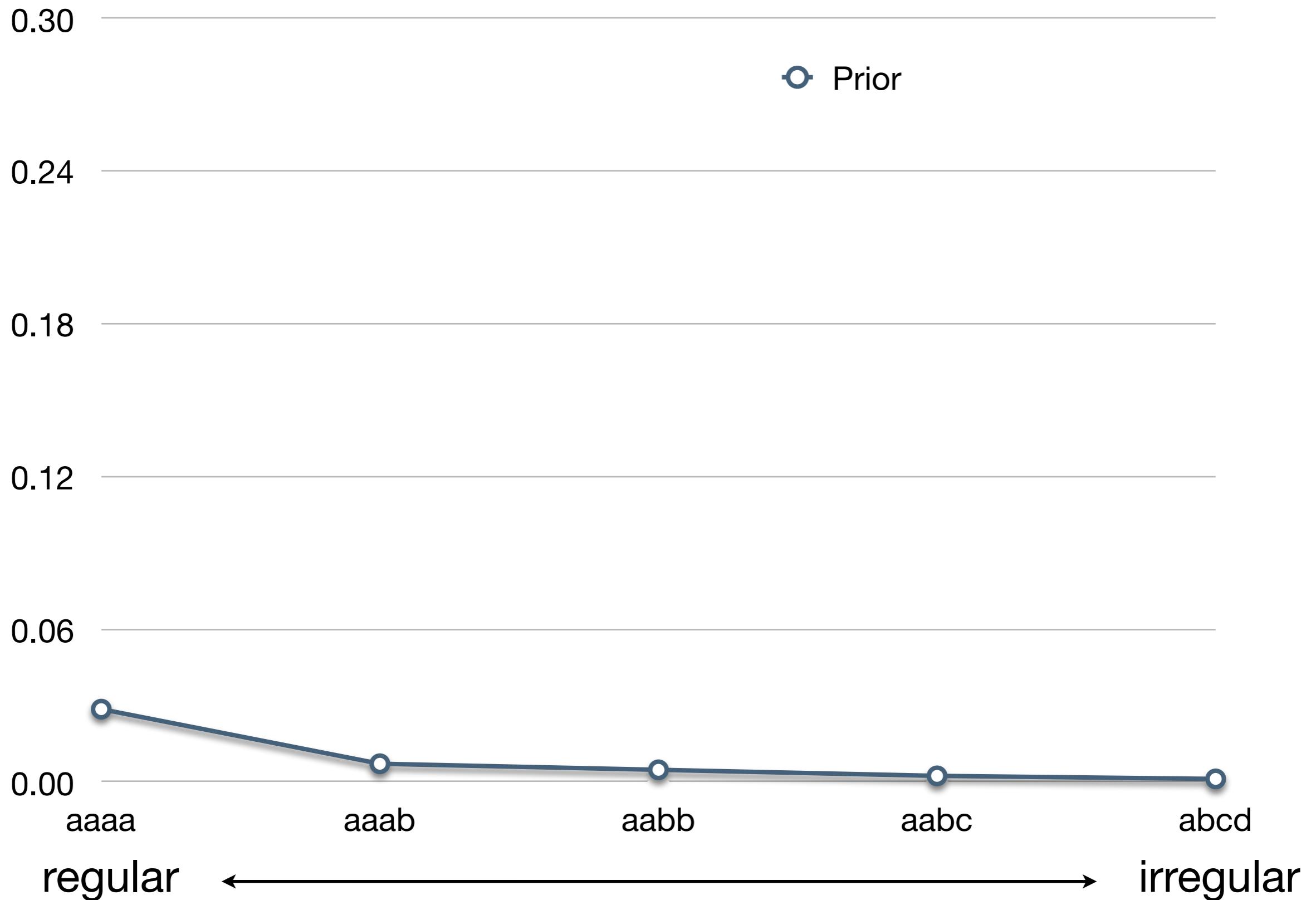
---

- Model language as a set of meanings
- These meanings can be expressed regularly, or irregularly
- Start with the assumption that there is a slight innate bias in favour of regularity
  - We can vary the strength of this bias
  - It is reasonable to assume a simple bias like this is not language-specific
- Assume learners pick the best (i.e. MAP) hypothesis. What happens?

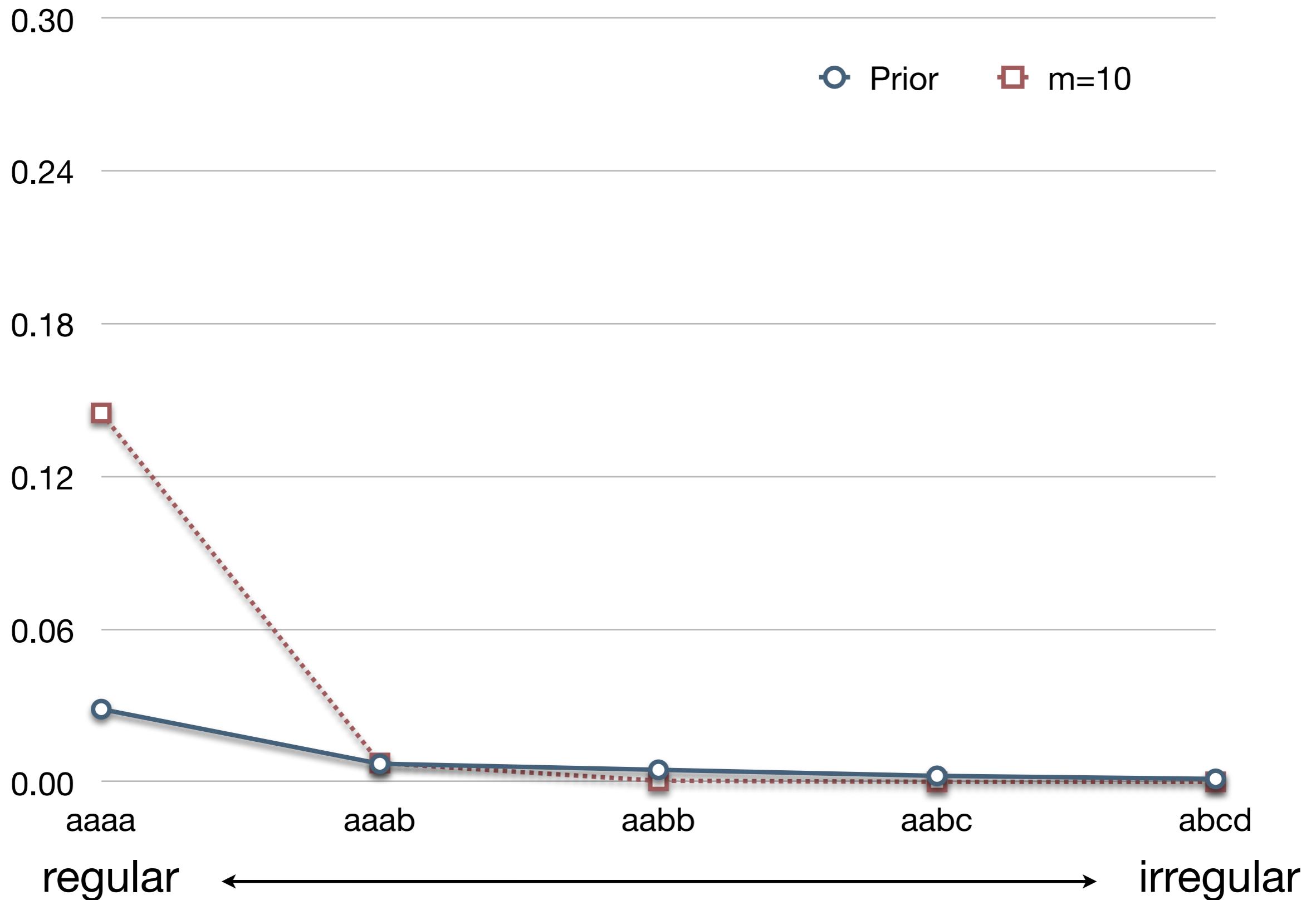
Probability of language by type: strong bias  
( $\alpha=1$ ,  $\varepsilon=0.05$ , 4 meanings, 4 classes)



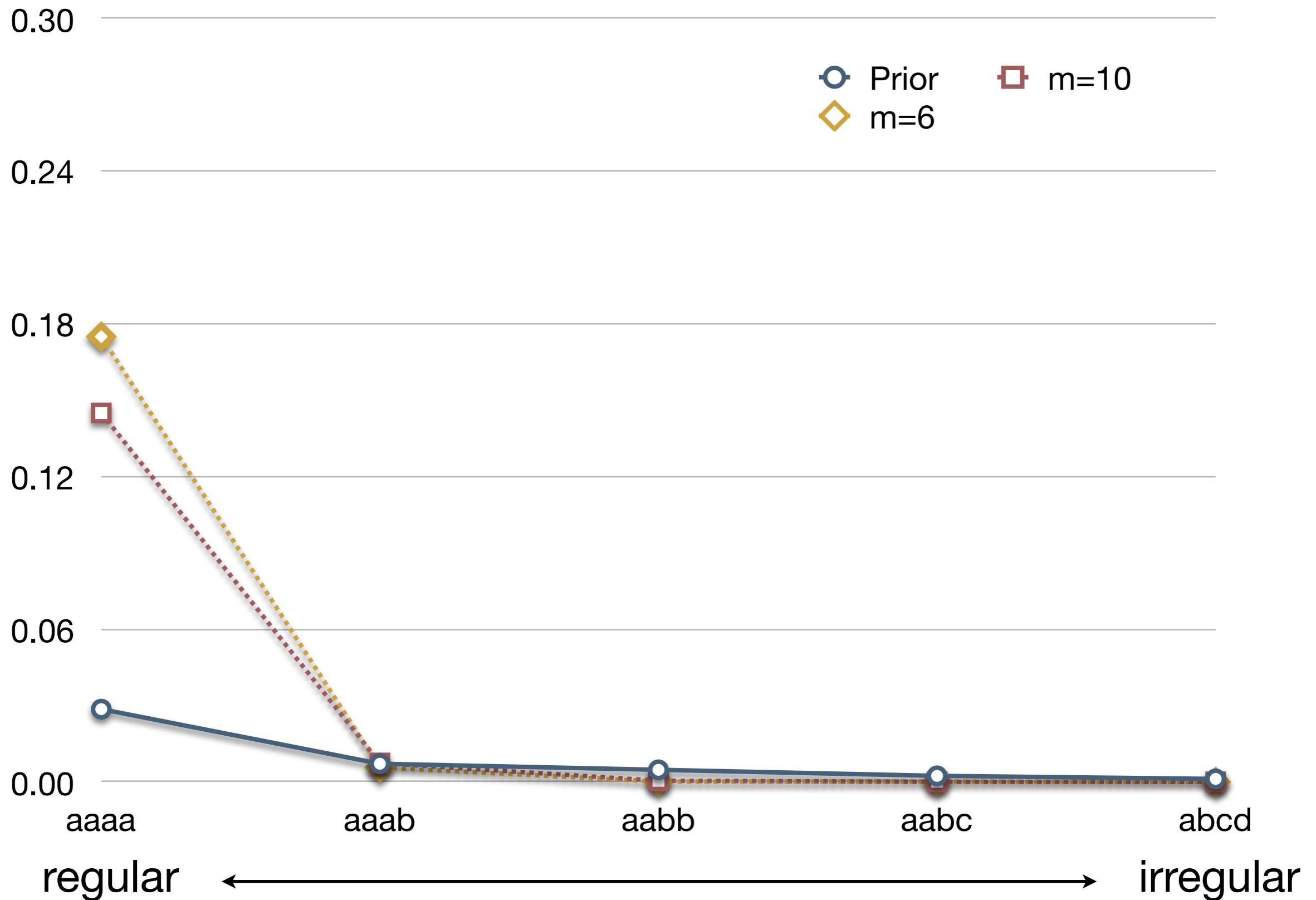
Probability of language by type: strong bias  
( $\alpha=1$ ,  $\varepsilon=0.05$ , 4 meanings, 4 classes)



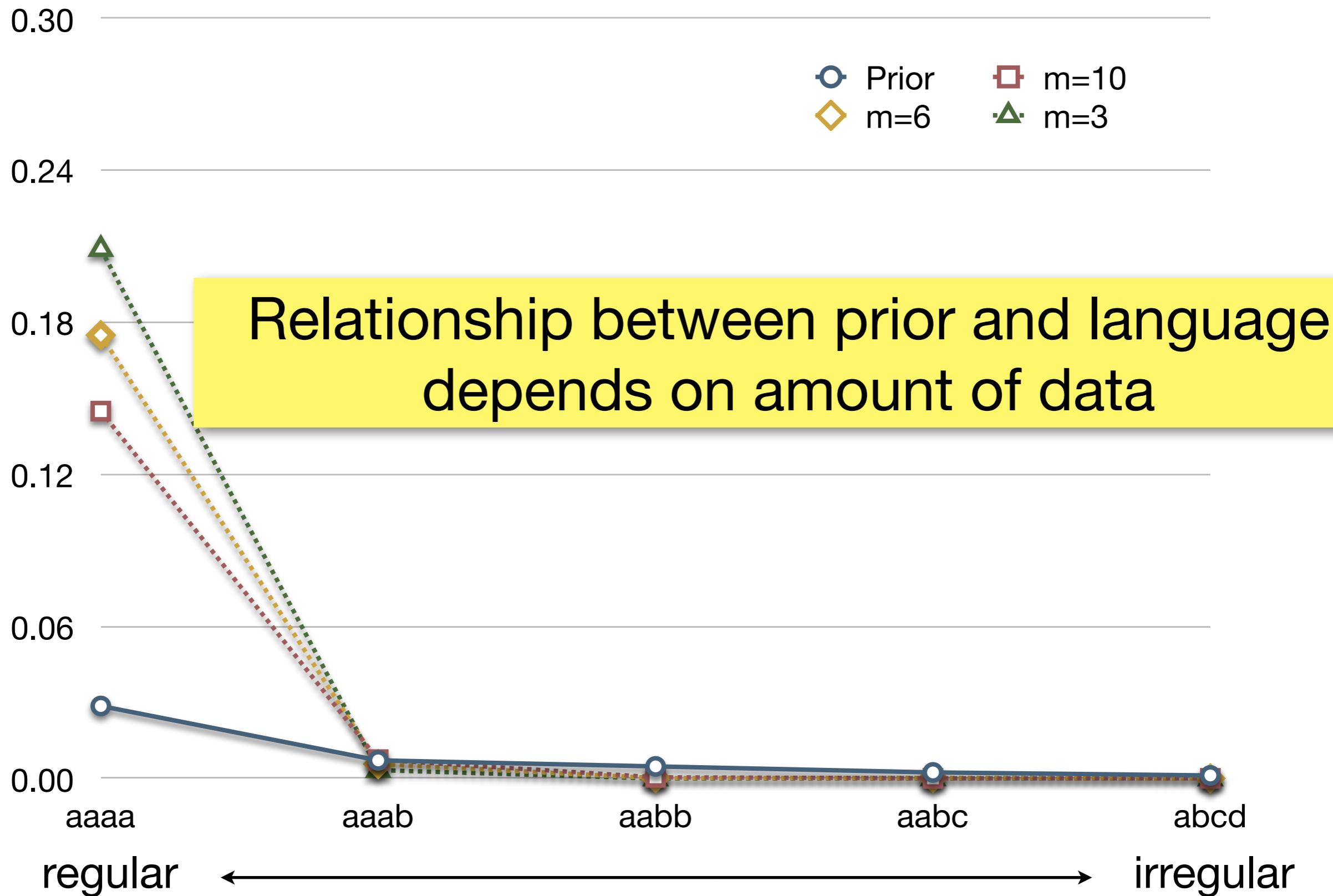
Probability of language by type: strong bias  
( $\alpha=1$ ,  $\varepsilon=0.05$ , 4 meanings, 4 classes)



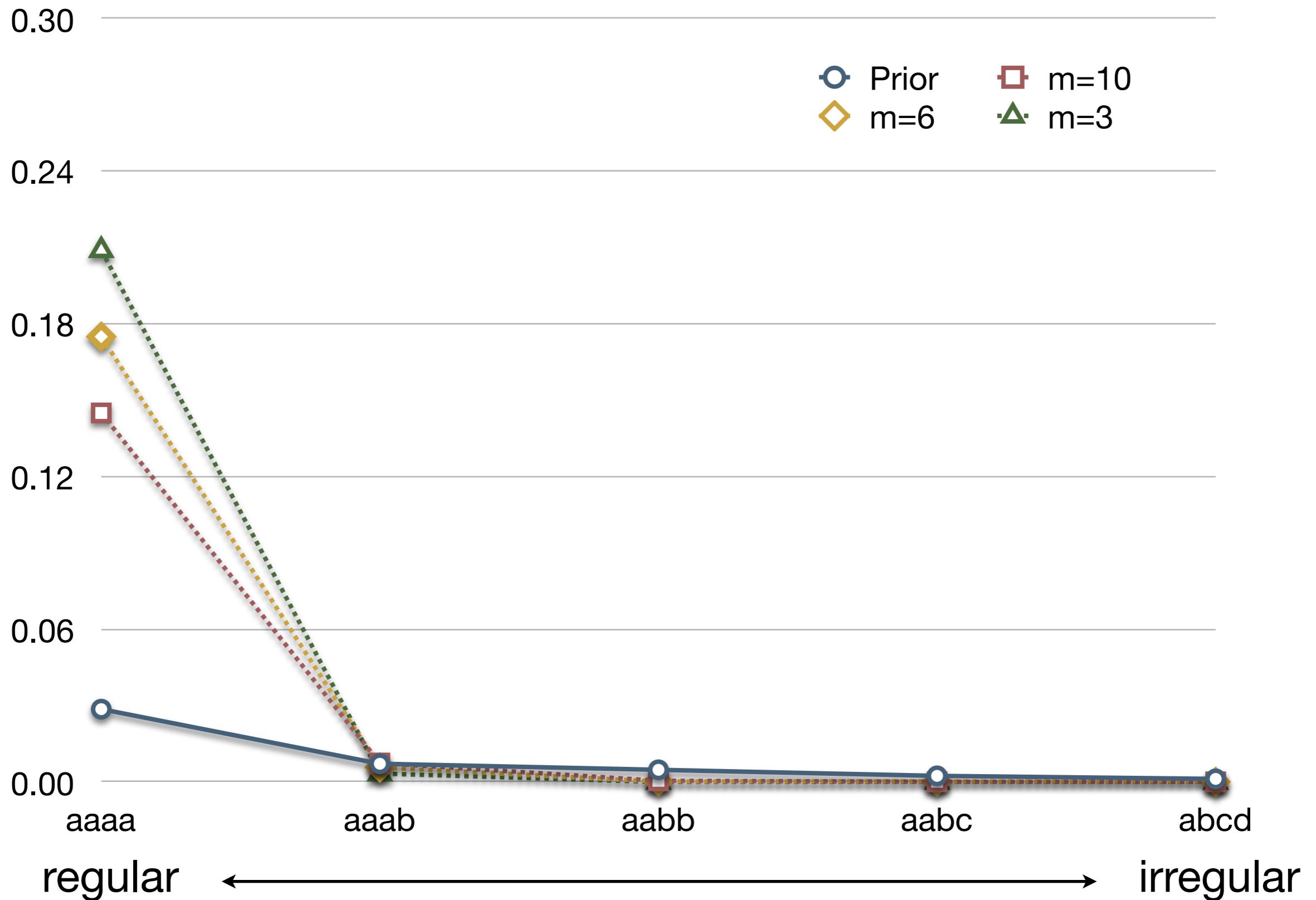
Probability of language by type: strong bias  
( $\alpha=1$ ,  $\varepsilon=0.05$ , 4 meanings, 4 classes)



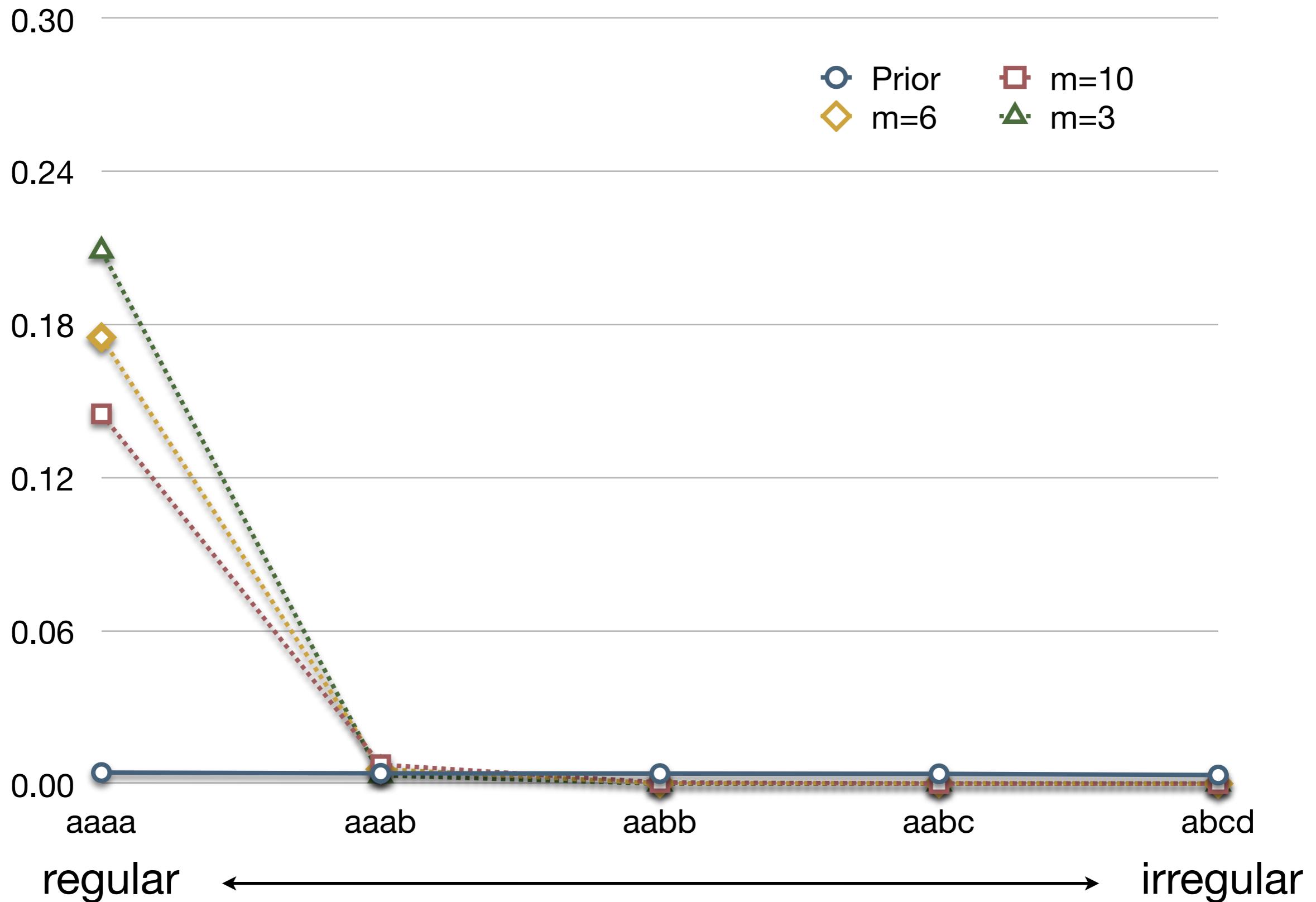
Probability of language by type: strong bias  
( $\alpha=1$ ,  $\varepsilon=0.05$ , 4 meanings, 4 classes)



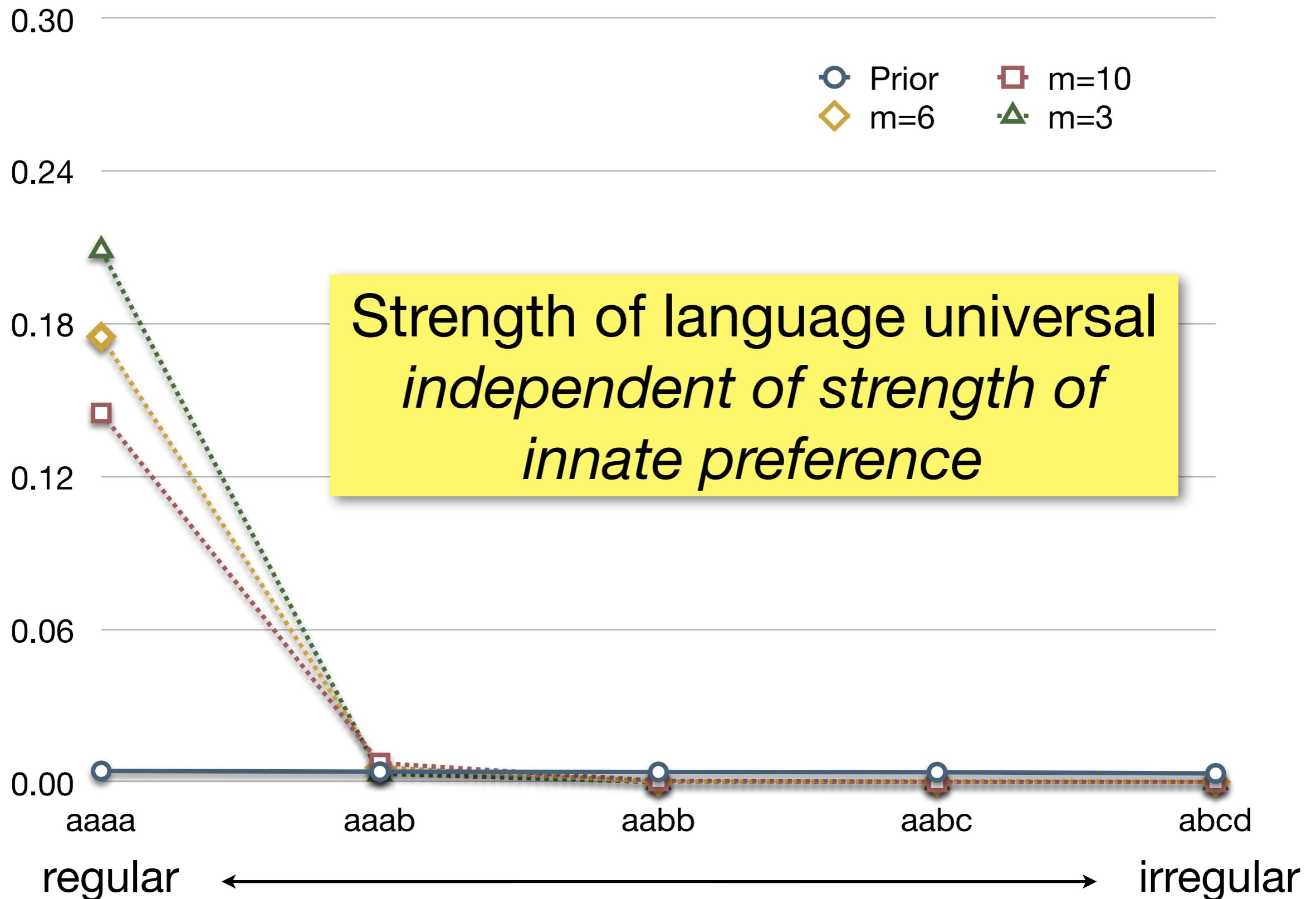
Probability of language by type: strong bias  
( $\alpha=1$ ,  $\varepsilon=0.05$ , 4 meanings, 4 classes)



Probability of language by type: weak bias  
( $\alpha=40$ ,  $\varepsilon=0.05$ , 4 meanings, 4 classes)



Probability of language by type: weak bias  
( $\alpha=40$ ,  $\varepsilon=0.05$ , 4 meanings, 4 classes)



# Conclusions

---

- Iterated Bayesian Learning allows us to more precisely understand the relationship between learning bias and eventual language structure
- If you assume social learning is about maximising the chance of converging on what other people are doing (i.e. selecting the MAP hypothesis), then cultural evolution does a lot of work for you
- Very weak innate biases are all that's needed to explain strong linguistic universals
- If we see universals in language, then we should not be assuming that these are hard-coded as strong constraints in the genes

# Extra material (if there is time)

---

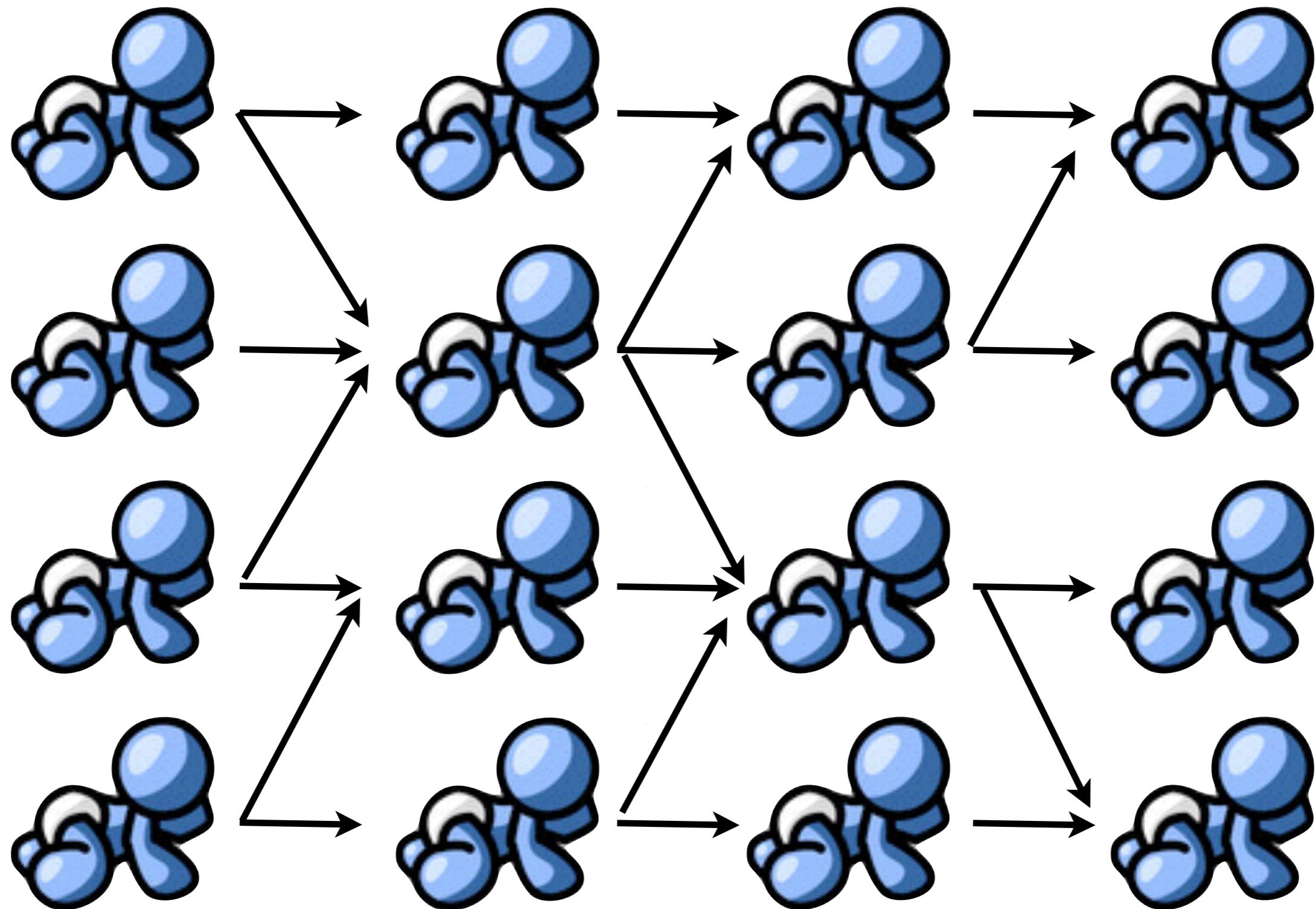
# Sampling vs MAP: which is right?

---

- If language learning is like sampling, language universals probably closely reflect learner biases. If it's like MAP, they don't.
- How can we tell which is right?
  - Run experiments on real people to see if they behave like they are sampling or selecting the MAP language
  - Maybe evolution will favour one alternative over the other?
    - See penultimate lecture
    - Maybe one of these results is an unrepresentative special case
      - For instance: what happens if we go beyond long skinny diffusion chains and look at transmission in populations?
      - Smith (2009), Burkett & Griffiths (2010)

# Moving to populations

---



# Sampler **populations** look like MAP populations

---

- In populations, when samplers learn from multiple teachers:

No convergence to the prior

Amplification of weak biases

Bottleneck effects

...

- **Play with this yourself in Monday's lab**

# References

---

- Griffiths, T. L. and Kalish, M. L. (2007) Language evolution by iterated learning with Bayesian agents. *Cognitive Science*, 31, 441-480.
- Kirby, S., Dowman, M. and Griffiths, T. (2007) Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences*, 104, 5241-5245.
- Smith, K. (2009). Iterated learning in populations of Bayesian agents. In N.A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31th Annual Conference of the Cognitive Science Society* (pp. 697-702). Austin, TX: Cognitive Science Society.