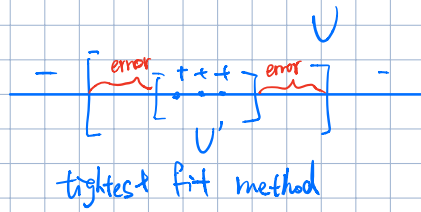


knowledge: has to be testable prediction.

Not all examples contribute the same amount of information.

per Occam's razor, simpler models are better



U' has an error rate of at most ϵ with probability p

errors $< \epsilon$.