Regularization & Bayesian Estimation

Mikio Aoi & Gal Mishne

UCSD NGP bootcamp, September 2020

Combining information from different sources

Many applications in neuroscience

- Want to make judgments about things we don’t see from things we do see.
- Integration across sensory systems (eg. normative models)
- Adding data / features we didn’t collect ourselves

Need a principled way to combine disparate pieces of information to a common purpose, vis a vis different levels of confidence in each piece.

Joint probabilities

unobserved parts
(model parameters, synaptic weights,
gene expression)

$$P(Y, X)$$

observed parts
(spikes, flourescence, voltage, etc.)

Conditional probabilities

$$\begin{aligned} P(Y,X) &= P(X|Y)P(Y) \\ &= P(Y|X)P(X) \end{aligned}$$

Conditional probabilities

$$P(X|Y)P(Y) = P(Y|X)P(X)$$

Bayes' rule

$$\overbrace{\boxed{P(X|Y)}}^{\text{posterior}} = \frac{\overbrace{\boxed{P(Y|X)}}^{\text{likelihood}}\ \overbrace{\boxed{P(X)}}^{\text{prior}}}{\underbrace{\boxed{P(Y)}}_{\text{marginal likelihood / evidence / partition function}}}$$

Bayes' rule

What are the chances you have covid if you tested negative?

$$P(\text{have}|-) = ?$$

Bayes' rule

What are the chances you have covid if you tested negative?

$$\underbrace{\boxed{P(X|Y)}}_{\text{posterior}} = \frac{\overbrace{\boxed{P(Y|X)}}^{\text{likelihood}}\ \overbrace{\boxed{P(X)}}^{\text{prior}}}{\underbrace{\boxed{P(Y)}}_{\text{marginal likelihood / evidence / partition function}}}$$

Bayes' rule

What are the chances you have covid if you tested negative?

$$\underbrace{\boxed{P(\text{have}|-)}}_{\text{posterior}} = \frac{\underbrace{\boxed{P(Y|X)}}_{\text{likelihood}}\ \underbrace{\boxed{P(X)}}_{\text{prior}}}{\underbrace{\boxed{P(Y)}}_{\text{marginal likelihood / evidence / partition function}}}$$

Bayes' rule

What are the chances you have covid if you tested negative?

$$\underbrace{\boxed{P(\text{have}|-)}}_{}^{\text{posterior}} = \frac{\overbrace{\boxed{P(-|\text{have})}}^{\text{false negative rate}}\;\overbrace{\boxed{P(X)}}^{\text{prior}}}{\underbrace{\boxed{P(Y)}}_{\text{marginal likelihood / evidence / partition function}}}$$

Bayes' rule

What are the chances you have covid if you tested negative?

false
negative rate

prevalence

posterior

$$\boxed{P(\text{have}|-)} = \frac{\boxed{P(-|\text{have})}\boxed{P(\text{have})}}{\boxed{P(Y)}}$$

marginal likelihood / evidence / partition function

Bayes' rule

What are the chances you have covid if you tested negative?

$$\underbrace{\boxed{P(\text{have}|-)}}_{\text{posterior}} = \frac{\overbrace{\boxed{P(-|\text{have})}}^{\text{false negative rate}}\ \overbrace{\boxed{P(\text{have})}}^{\text{prevalence}}}{\underbrace{\boxed{P(-)}}_{\text{negativity rate}}}$$

Bayes' rule

What are the chances you have covid if you tested negative?

90% CI: (.02,.54)

prevalence

posterior

$$\boxed{P(\text{have}|-)} = \frac{\boxed{P(-|\text{have})}\,\boxed{P(\text{have})}}{\boxed{P(-)}}$$

negativity rate

https://www.medrxiv.org/content/10.1101/2020.04.16.20066787v2

Bayes' rule

What are the chances you have covid if you tested negative?

90% CI:
(.02,.54)

posterior

$$\boxed{P(\text{have}|-)} = \frac{\boxed{P(-|\text{have})}\boxed{P(\text{have})}}{\boxed{P(-)}}$$

varies a lot by geographic location, occupation, etc.

Sampling and testing biases

negativity rate

Bayes' rule

What are the chances you have covid if you tested negative?

90% CI: (.02,.54)

0.01 for closest studied region (San Francisco)

posterior

$$\boxed{P(\text{have}|-)} = \frac{\boxed{P(-|\text{have})}\,\boxed{P(\text{have})}}{\boxed{P(-)}}$$

negativity rate

https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/commercial-lab-surveys.html

Bayes' rule

What are the chances you have covid if you tested negative?

90% CI: (.02,.54)

0.01 for closest studied region (San Francisco)

posterior

$$\boxed{P(\text{have}|-)} = \frac{\boxed{P(-|\text{have})}\boxed{P(\text{have})}}{\boxed{P(-)}}$$

$$\boxed{P(-) = P(-|\text{have})P(\text{have}) + P(-|\text{don't have})P(\text{don't have})}$$

false positive prevalence

true negative =1-false positive

1-prevelence

false positives <0.05

Bayes' rule

What are the chances you have covid if you tested negative?

90% CI: (.02,.54)

0.01 for closest studied region (San Francisco)

posterior

$$\boxed{P(\text{have}|-)} = \frac{\boxed{P(-|\text{have})}\,\boxed{P(\text{have})}}{\boxed{P(-)}}$$

$$\boxed{P(-) = P(-|\text{have})P(\text{have}) + P(-|\text{don't have})P(\text{don't have})}$$

(.2,.54) 0.01 true negative =1- false positive .99

false positives <0.05

Bayes' rule

What are the chances you have covid if you tested negative?

posterior

$$\boxed{P(\text{have}|-)} = \frac{(.02,.54) * 0.01}{(.02,.54) * 0.01 + 0.95 * .99}$$

$$= (.0002, .0057)$$

Bayes' rule

What are the chances you have covid if you tested negative?

posterior

$$\boxed{P(\text{have} \mid -)} = \frac{(.02,.54) * 0.01}{.991 \text{ for UCSD students}}$$

$$= (.0002, .0054)$$

Bayes' rule

What are the chances you have covid if you tested negative?

posterior

$$\boxed{P(\text{have} \mid -)} = \frac{(.02,.54) * .23}{(.02,.54) * .23 + 0.95 * .77}$$

$$= \quad (.0062,\ .1451)$$

NYC is a different story

Bayes' rule with Gaussians

Tutorial scenario

we get a sample of these

$$Y_1 \sim \mathcal{N}(x, \sigma_1^2) \qquad Y_2 \sim \mathcal{N}(x, \sigma_2^2)$$

want to know *x*

Bayes' rule with Gaussians

Point estimation

we get a sample of this

$$Y_1 \sim \mathcal{N}(x, \sigma_1^2) \qquad X \sim \mathcal{N}(\theta, \tau^2)$$

$$p(Y|X)$$

want to know *x*

Bayes' rule with Gaussians

Point estimation

likelihood

$$p(Y|X)$$

prior

$$p(X)$$

$$Y_1 \sim \mathcal{N}(x, \sigma_1^2) \qquad X \sim \mathcal{N}(\theta, \tau^2)$$

want to know *x*

Bayes' rule with Gaussians

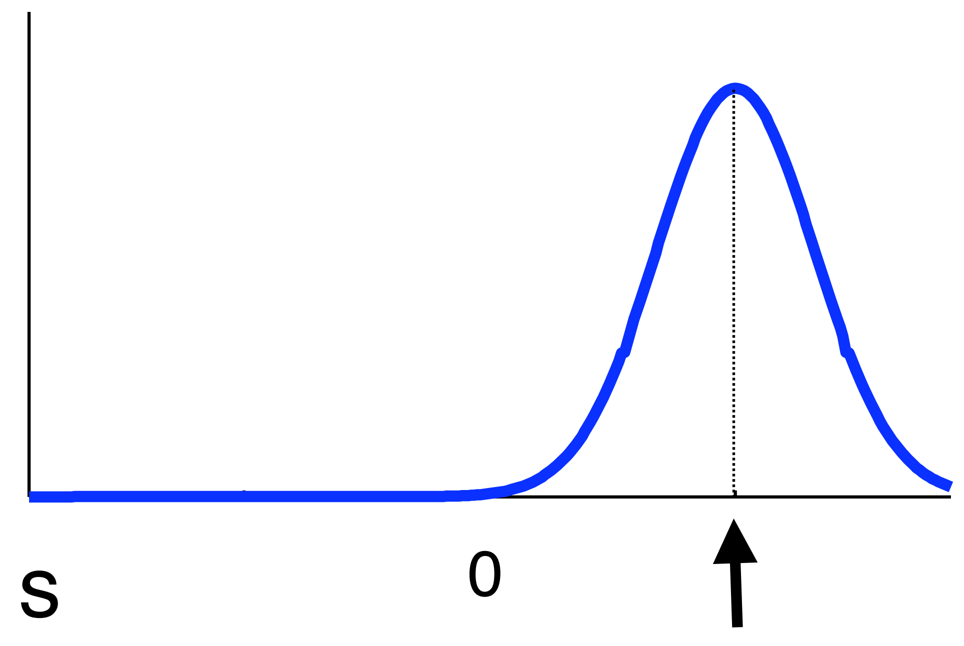
Bayesian estimates are biased

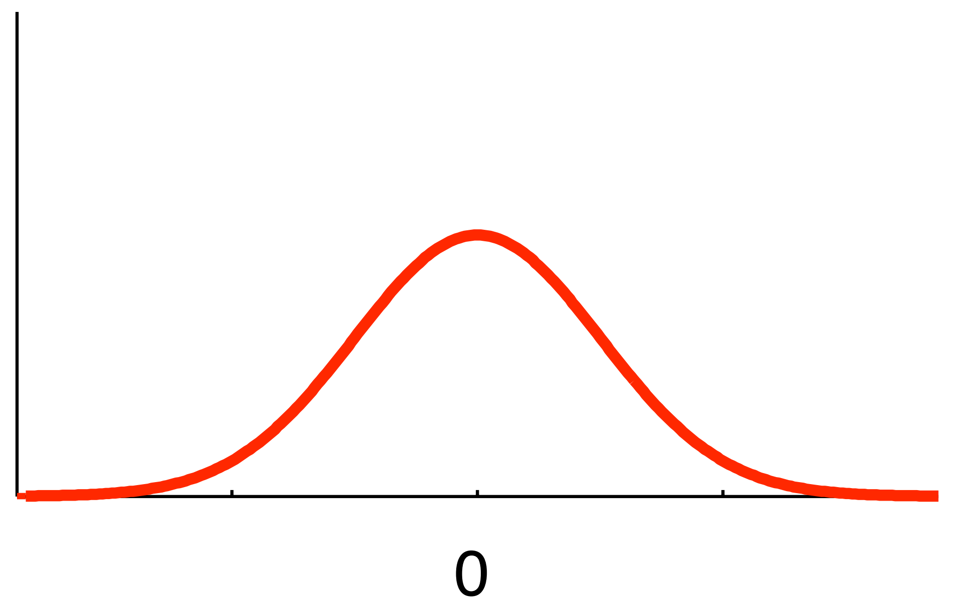
likelihood

$$p(Y|X)$$

prior

$$p(X)$$



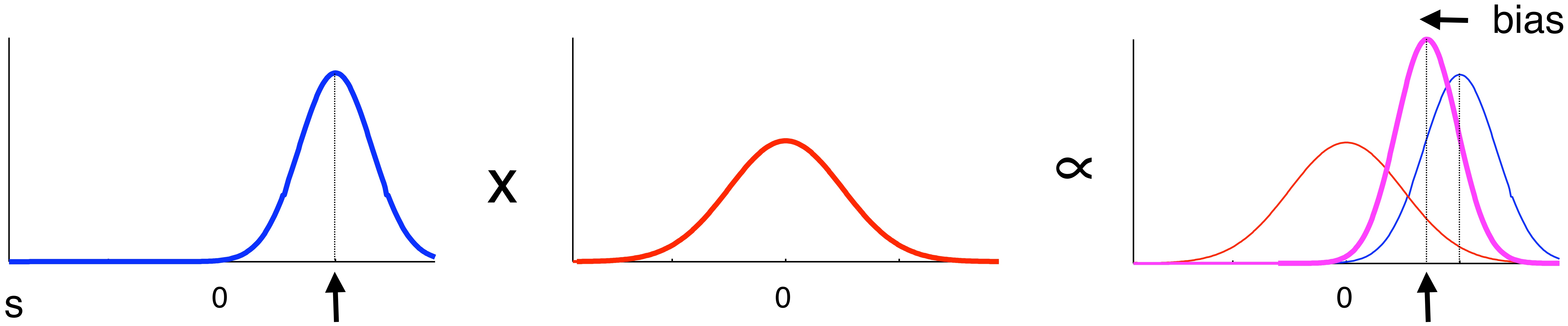


$$\hat{X}_{MLE}$$

Bayes' rule with Gaussians

Bayesian estimates are biased

likelihood $p(Y|X)$ prior $p(X)$ posterior $p(X|Y)$



$\hat{X}_{MLE}$ $\hat{X}_{MAP}$

- MAP estimate lives between ML estimate & prior

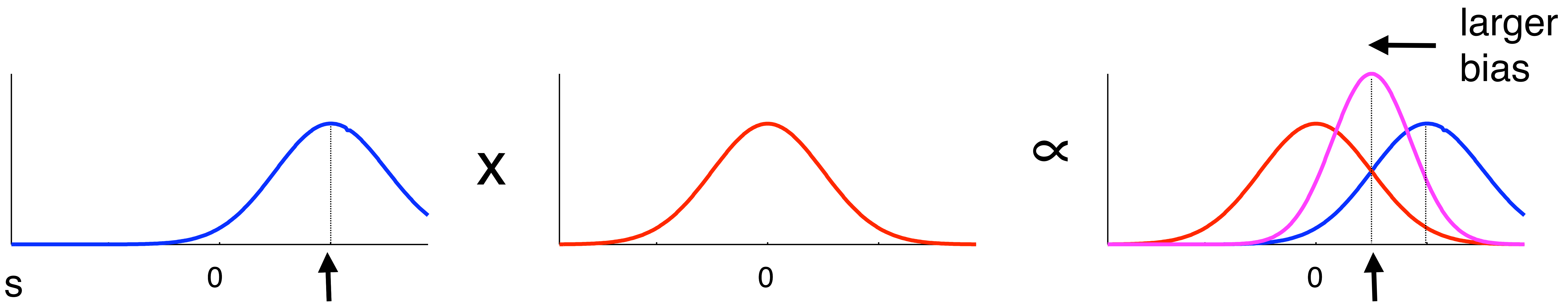
Bayes' rule with Gaussians

Bayesian estimates are biased

likelihood $p(Y|X)$

prior $p(X)$

posterior $p(X|Y)$

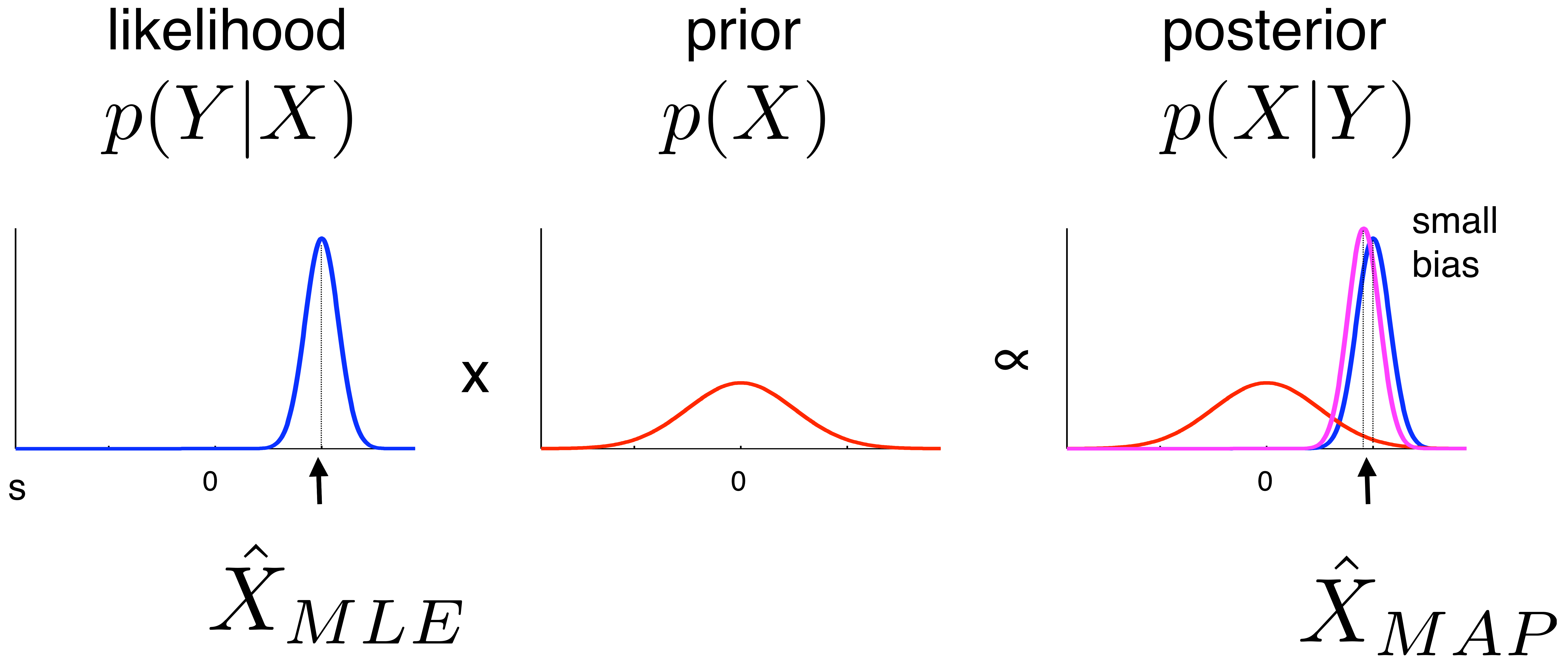


$\hat{X}_{MLE}$

$\hat{X}_{MAP}$

Bayes' rule with Gaussians

Bayesian estimates are biased



Bayes' rule with Gaussians

Point estimation

likelihood

$$p(Y|X)$$

prior

$$p(X)$$

$$Y_1 \sim \mathcal{N}(x, \sigma_1^2) \qquad X \sim \mathcal{N}(\theta, \tau^2)$$

want to know *x*

Bayes' rule with Gaussians

Point estimation

likelihood

$$p(Y|X)$$

prior

$$p(X)$$

$$Y_n \sim \mathcal{N}(x, \sigma^2) \qquad X \sim \mathcal{N}(\theta, \tau^2)$$

want to know *x*

Bayes' rule with Gaussians

Point estimation

likelihood

$$p(Y|X)$$

prior

$$p(X)$$

$$Y_n \sim \mathcal{N}(x, \sigma^2)$$

$$X \sim \mathcal{N}(\theta, \tau^2)$$

maximum likelihood

$$\log(p(Y_1, Y_2, \ldots|X))$$

$$\hat{X}_{\mathrm{MLE}} = \arg\max_X \log(p(Y_1, Y_2, \ldots|X))$$

Bayes' rule with Gaussians

Point estimation

likelihood

$$p(Y|X)$$

prior

$$p(X)$$

$$Y_n \sim \mathcal{N}(x, \sigma^2)$$

$$X \sim \mathcal{N}(\theta, \tau^2)$$

maximum likelihood

$$\log(p(Y_1, Y_2, \ldots | X))$$

$$\hat{X}_{\mathrm{MLE}} = \arg\min_X \boxed{\frac{1}{N} \sum_n (X - Y_n)^2}$$

MSE

Bayes' rule with Gaussians

Regularized point estimation is Bayesian point estimation

likelihood

$$Y_n \sim \mathcal{N}(x, \sigma^2)$$

prior

$$X \sim \mathcal{N}(0, \tau^2)$$

$$\underbrace{\log(p(x|Y_1, Y_2, \ldots))}_{\text{log posterior}} \propto \underbrace{\frac{1}{N}\sum_n (X - Y_n)^2}_{\text{MSE}} + \underbrace{\beta X^2}_{\text{regularizer}}$$

"ridge", "ℓ_2"

penalty $\beta = \frac{\sigma^2}{\tau^2}$

NMA Tutorials

- Tutorial 1 - Regularization
  - Pick up where we left off with logistic regression
  - Regularizer is an expression of prior beliefs (in the Bayesian sense)
- Tutorial 2 - 4 - Combining information from 2 sources and manipulating the probabilities
- Tutorial 2 errata
  - typos in the instructional text - formula for mean of a distribution
  - url for videos are not all working. Will post correct url's in Slack.