Bayesian methods
Brain and cog perspectives

Mehrdad Jazayeri

June 9th, 46-3189, 4pm-6pm

MIT
Tutorials on advanced topics, 2015

An intuition for Bayesian estimation

<u>Medical testing</u>

A random person goes to the doctor to get a medical test for a rare disease. The test is pretty accurate: it gives a positive result for 99% of those who have the disease, and gives negative result for 90% of those who do not have the disease.

What are the chances that this person has the disease if the test comes out positive?

A. Less than 90%
B. 90%
C. between 90 and 99%
D. 99%
E. I don't know.

From Dan Levy's Master Class on teaching

$p(Yes)$ $p(+|Yes)$ $p(-|No)$

$p(Yes)$	$p(+\|Yes)$	$p(-\|No)$
2%	*99%*	*90%*

rare →

	$p(Yes)$	$p(+\|Yes)$	$p(-\|No)$
	2%	99%	90%
1000	20	$\sim 20^{+}/20^{+}$	$882^{-}/980^{-}$ $98^{+}/980^{-}$

$$p(Yes|+)$$

	$p(Yes)$	$p(+\mid Yes)$	$p(-\mid No)$
	2%	99%	90%
1000	20	~$20^{+}/20^{+}$	$882^{-}/980^{-}$ $98^{+}/980^{-}$

$$p(Yes|+)$$

20/(20+98) = 0.17

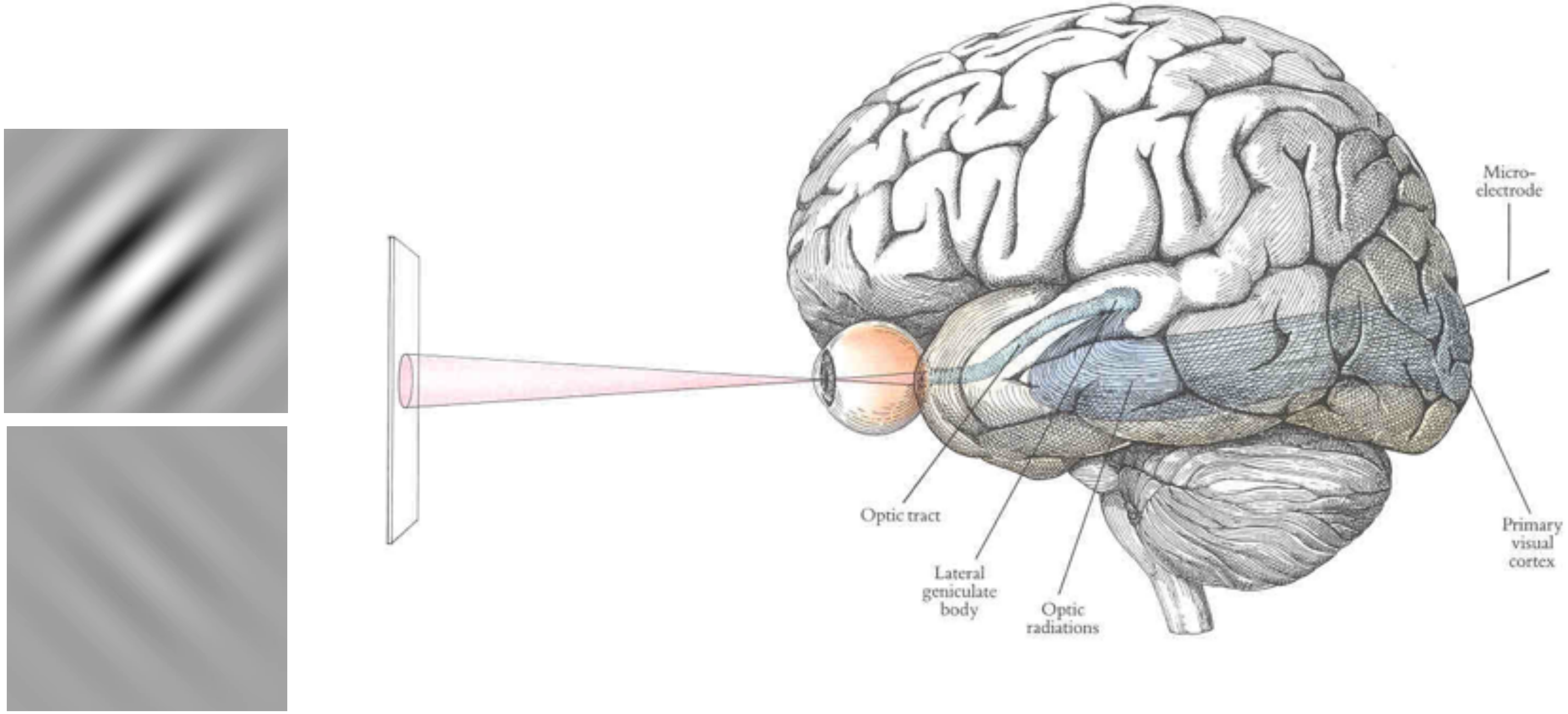
17%

$p(Yes)$	$p(+\|Yes)$	$p(-\|No)$
2%	99%	90%

$$p(Yes|+) = \frac{\overset{99\%}{p(+|Yes)}\overset{2\%}{p(Yes)}}{p(+)}$$

$$p(+) = \overset{99\%}{p(+|Yes)}\overset{2\%}{p(Yes)} + \overset{10\%}{p(+|No)}\overset{98\%}{p(No)}$$

Estimating visual contrast from neural activity



Estimating visual contrast from neural activity

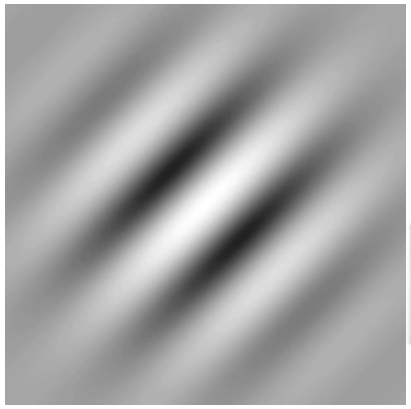
noise process *estimator*

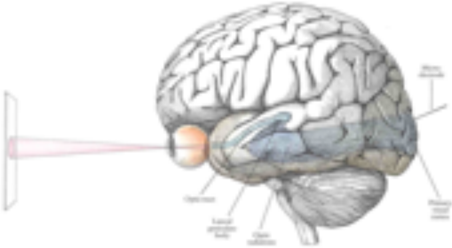
$$S \xrightarrow{p(m|S)} m^* \xrightarrow{S = f(m)} S^*$$

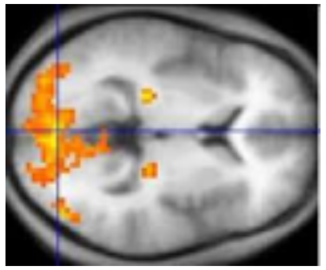
stimulus

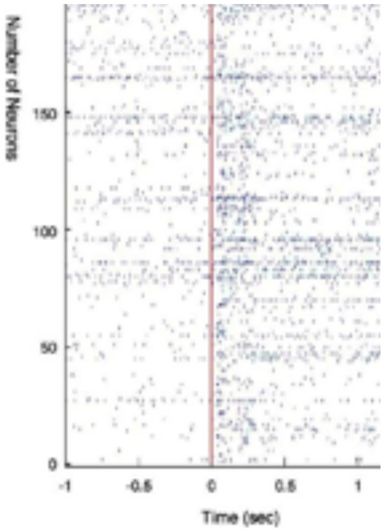
noisy measurement

estimate of the stimulus









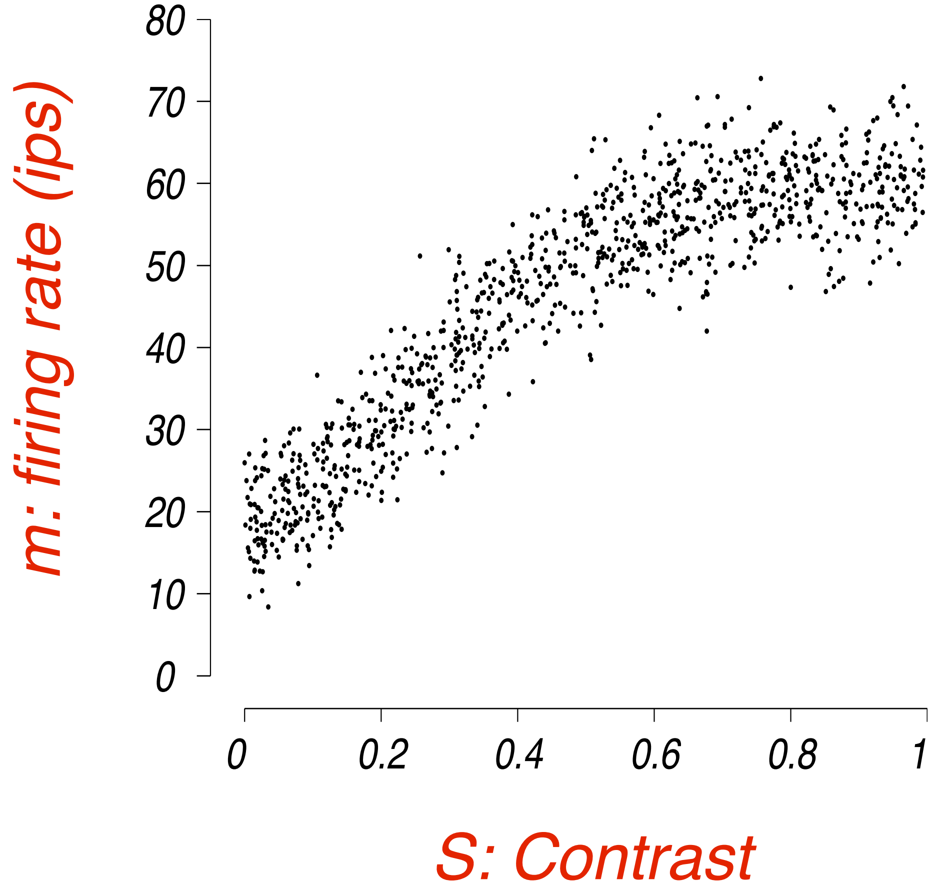
Conditional probability & likelihood function

noise process

$$S \xrightarrow{p(m|S)} m^*$$

stimulus

noisy measurement



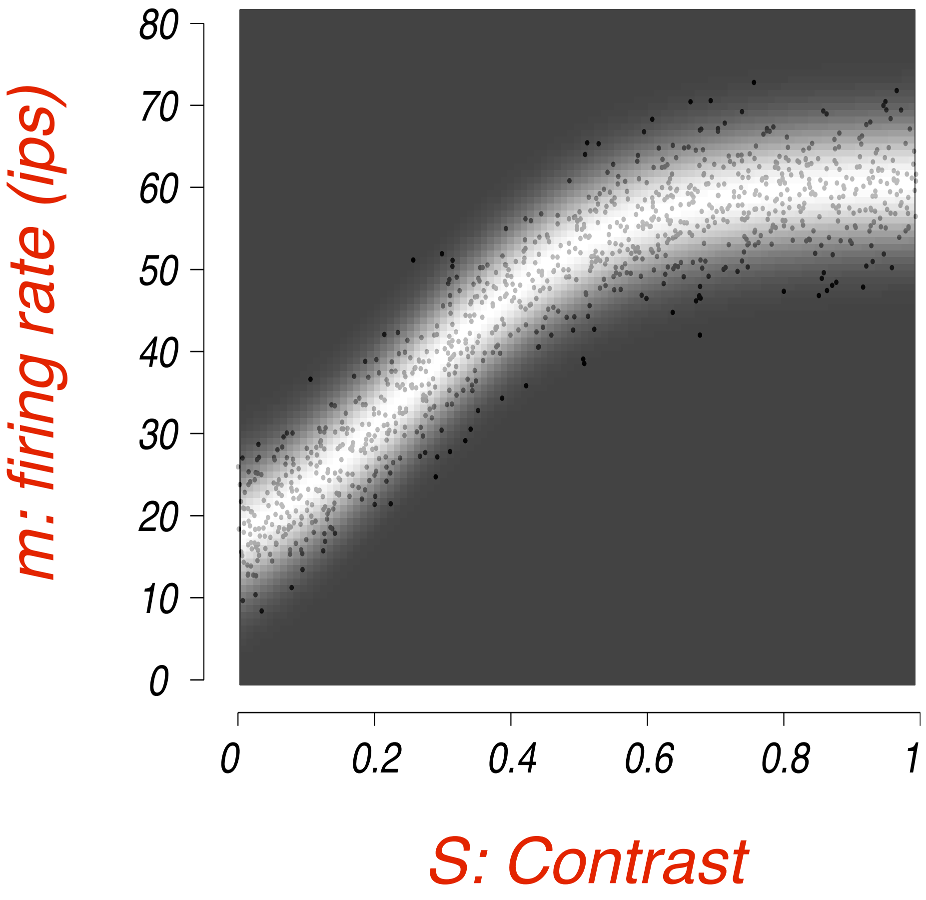
Conditional probability & likelihood function

noise process

$$S \xrightarrow{p(m|S)} m^*$$

stimulus

noisy measurement



Conditional probability & likelihood function

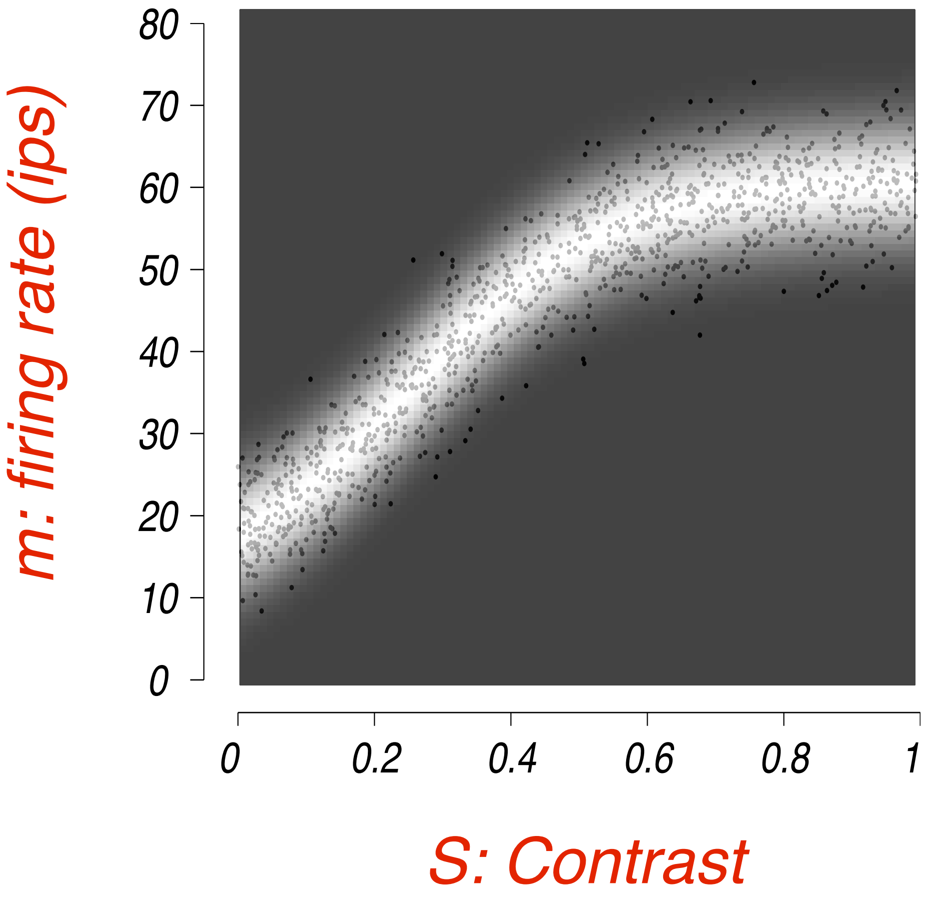
noise process

$$p(m|S)$$

$$S \xrightarrow{\hspace{3em}} m^*$$

stimulus

noisy measurement



Conditional probability

$$p(m|S)$$

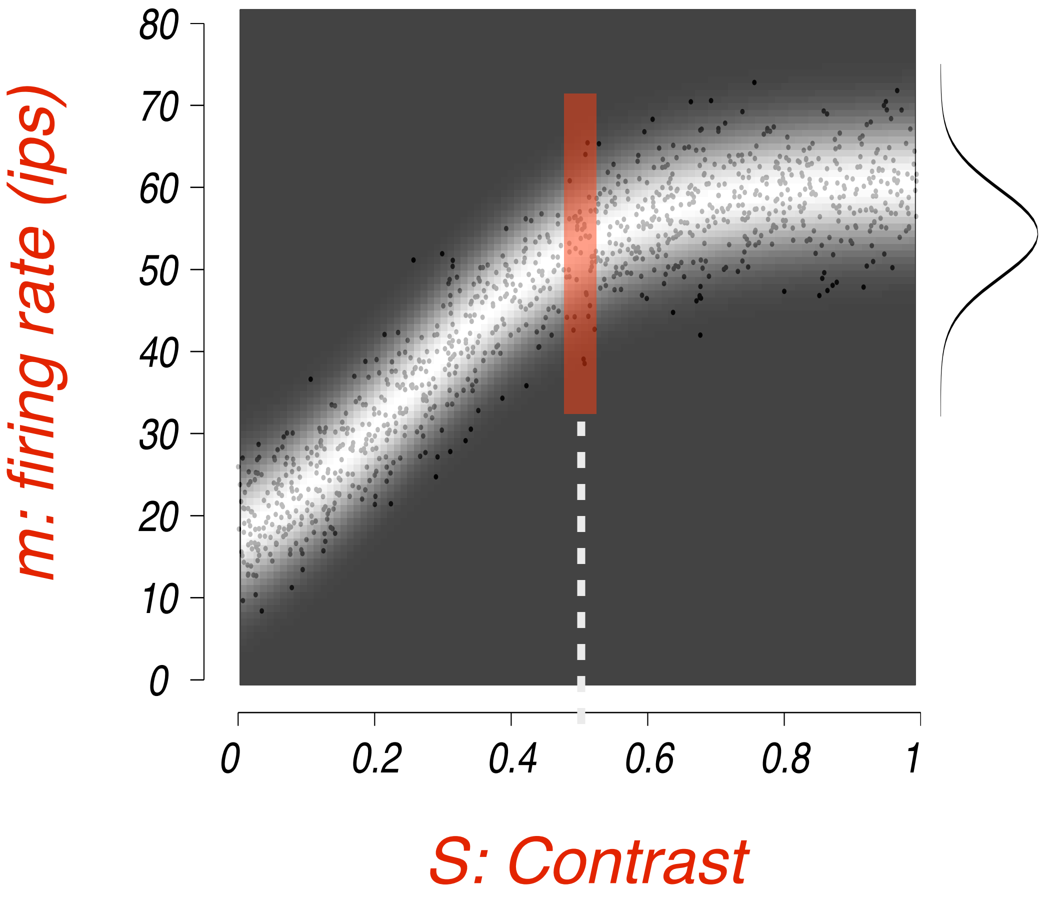
Conditional probability & likelihood function

noise process

$$S \xrightarrow{p(m|S)} m^*$$

stimulus

noisy measurement



Conditional probability

$$p(m|S)$$

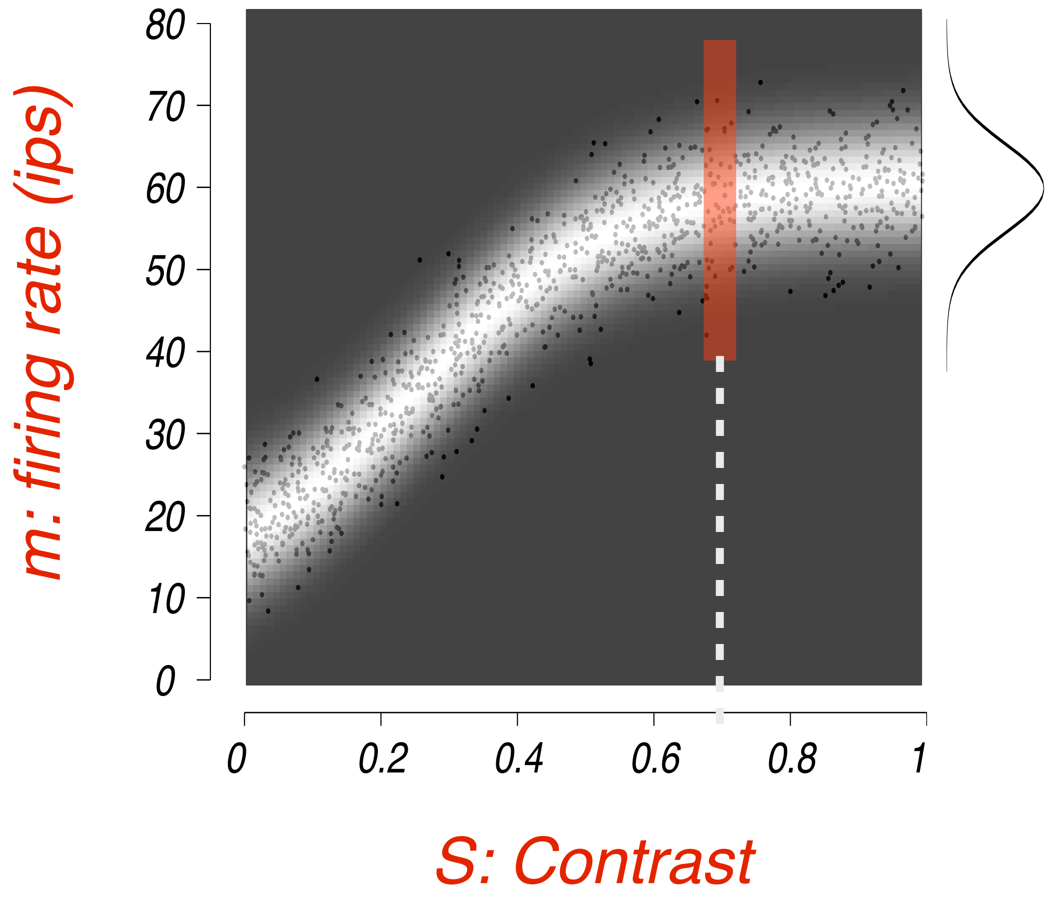
Conditional probability & likelihood function

noise process

$$S \xrightarrow{p(m|S)} m^*$$

stimulus

noisy measurement



Conditional probability

$$p(m|S)$$

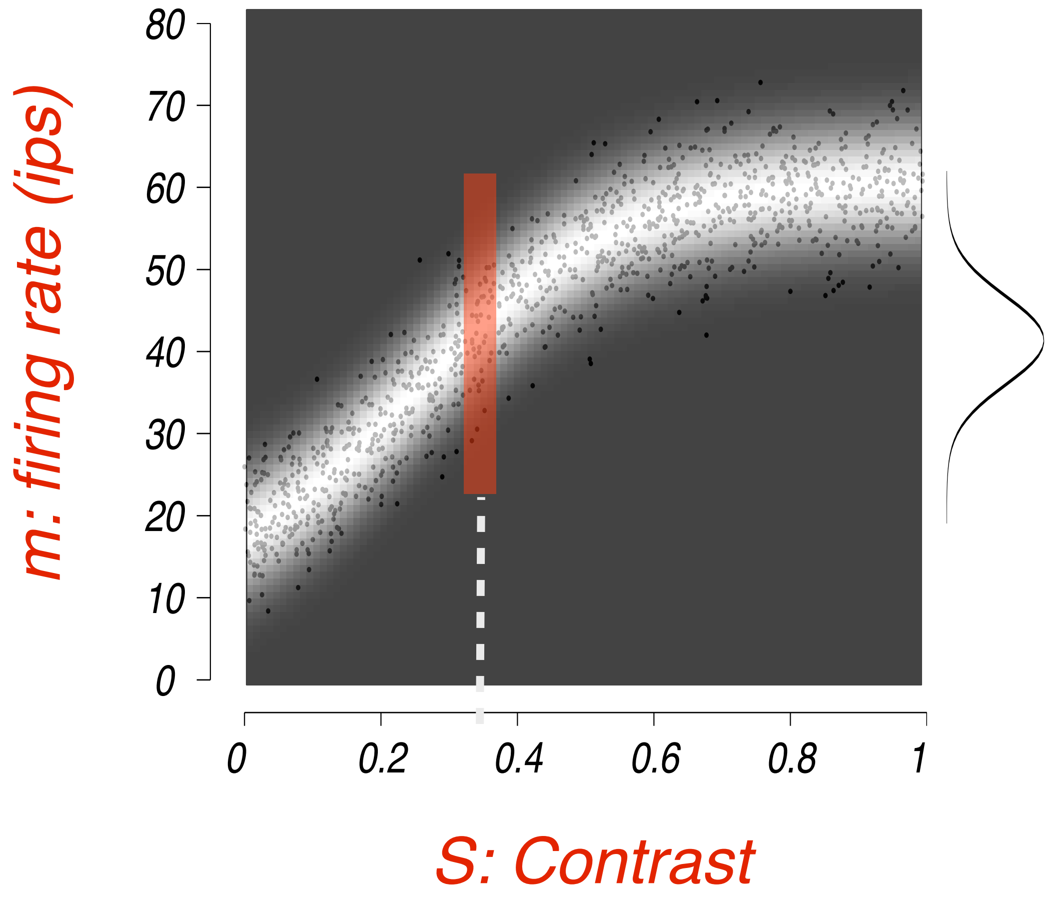
Conditional probability & likelihood function

noise process

$$S \xrightarrow{p(m|S)} m^*$$

stimulus

noisy measurement



Conditional probability

$$p(m|S)$$

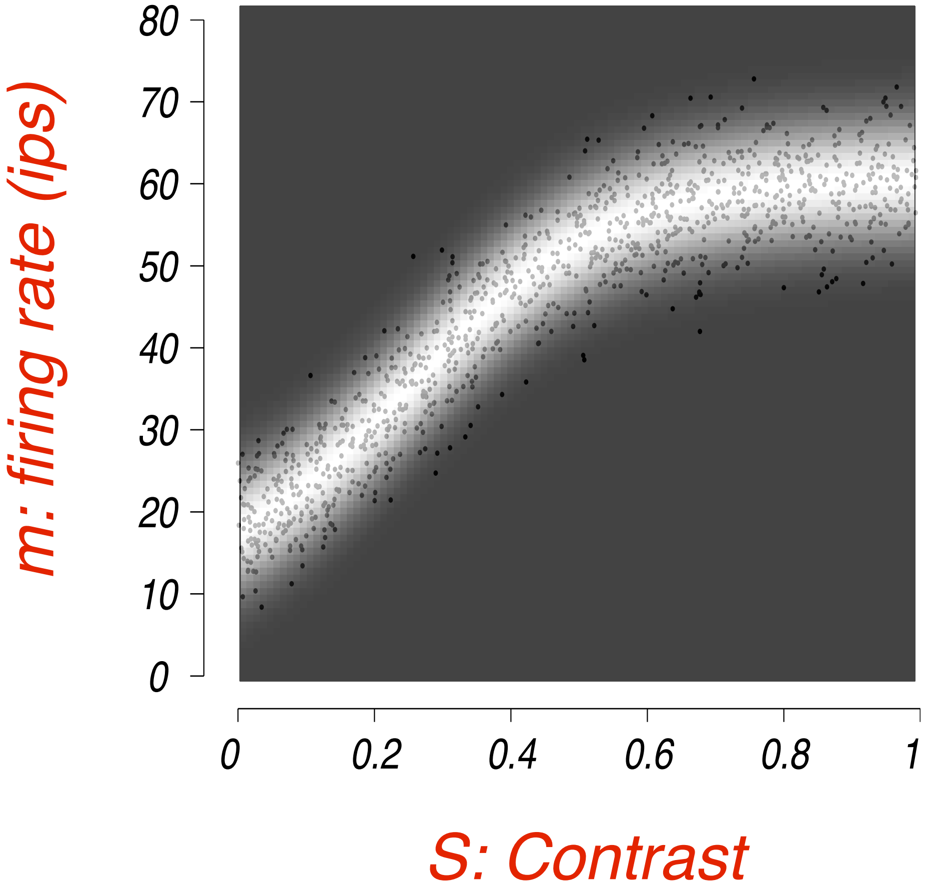
Conditional probability & likelihood function

noise process

$$S \xrightarrow{p(m|S)} m^*$$

stimulus

noisy measurement



Likelihood

$$L(S;m)$$

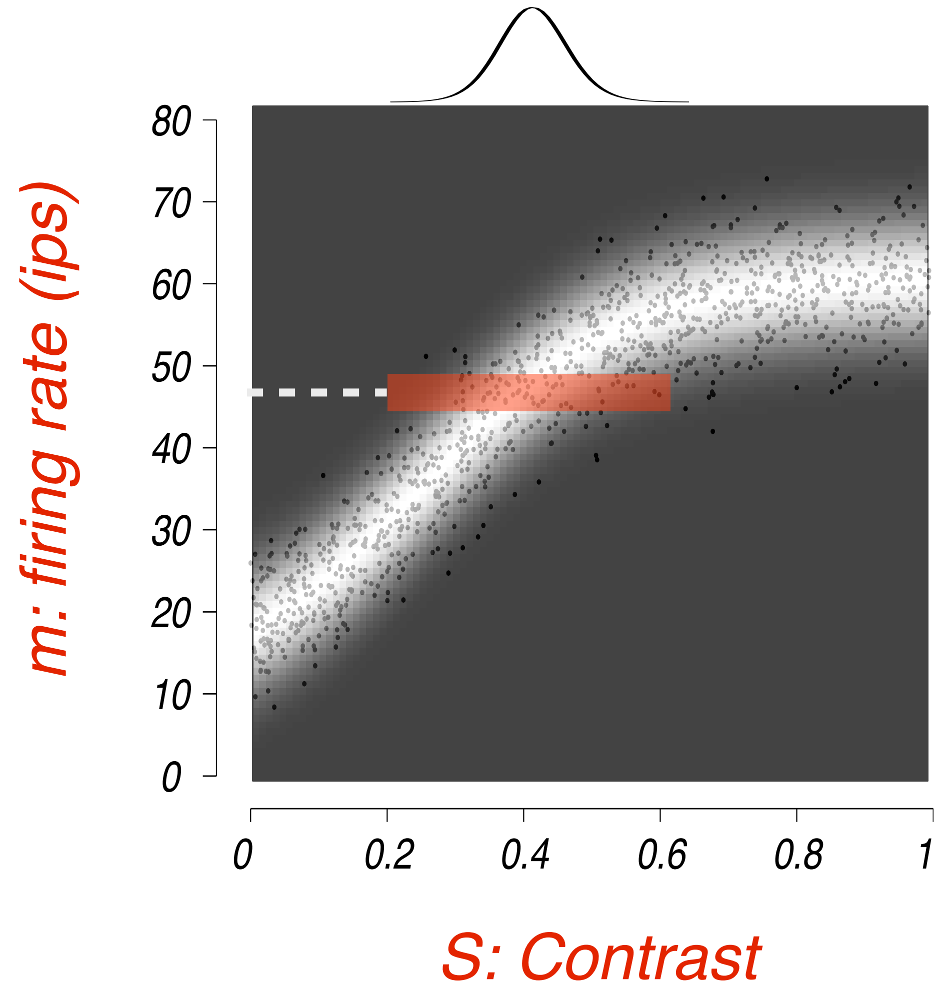
Conditional probability & likelihood function

noise process

$$S \xrightarrow{p(m|S)} m^*$$

stimulus

noisy measurement



Likelihood

$$L(S;m) = p(m|S)$$

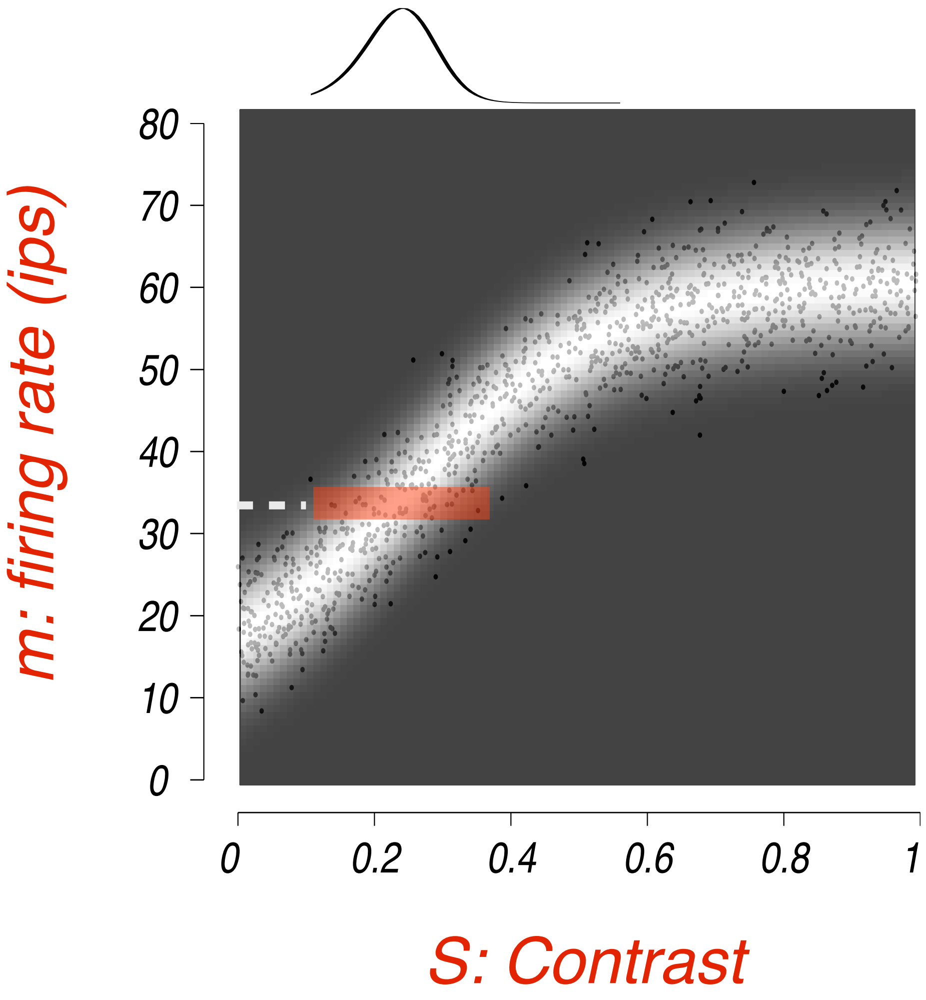
Conditional probability & likelihood function

noise process

$$S \xrightarrow{p(m|S)} m^*$$

stimulus

noisy measurement



Likelihood

$$L(S; m) = p(m|S)$$

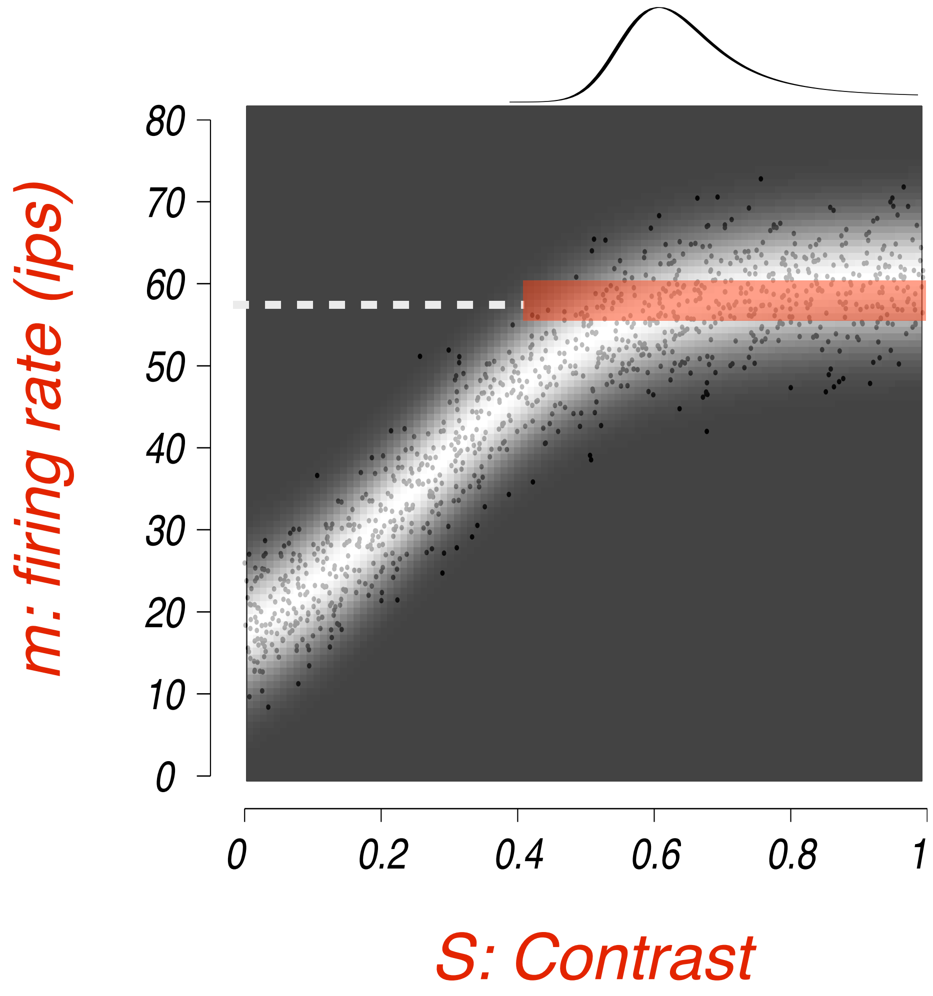
Conditional probability & likelihood function

noise process

$$S \xrightarrow{p(m|S)} m^*$$

stimulus

noisy measurement



Likelihood

$$L(S;m) = p(m|S)$$

Estimation

noise process

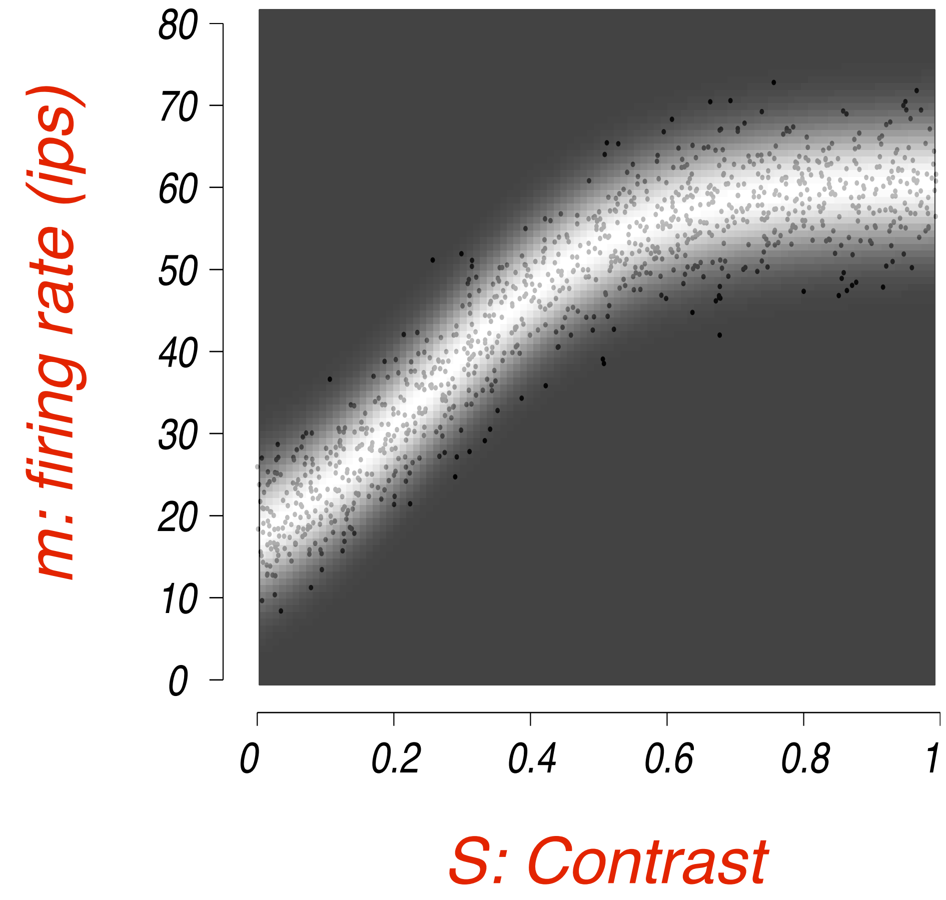
estimator

$$S \xrightarrow{p(m|S)} m^* \xrightarrow{S = f(m)} S^*$$

stimulus

noisy measurement

estimate of the stimulus



Estimation

noise process

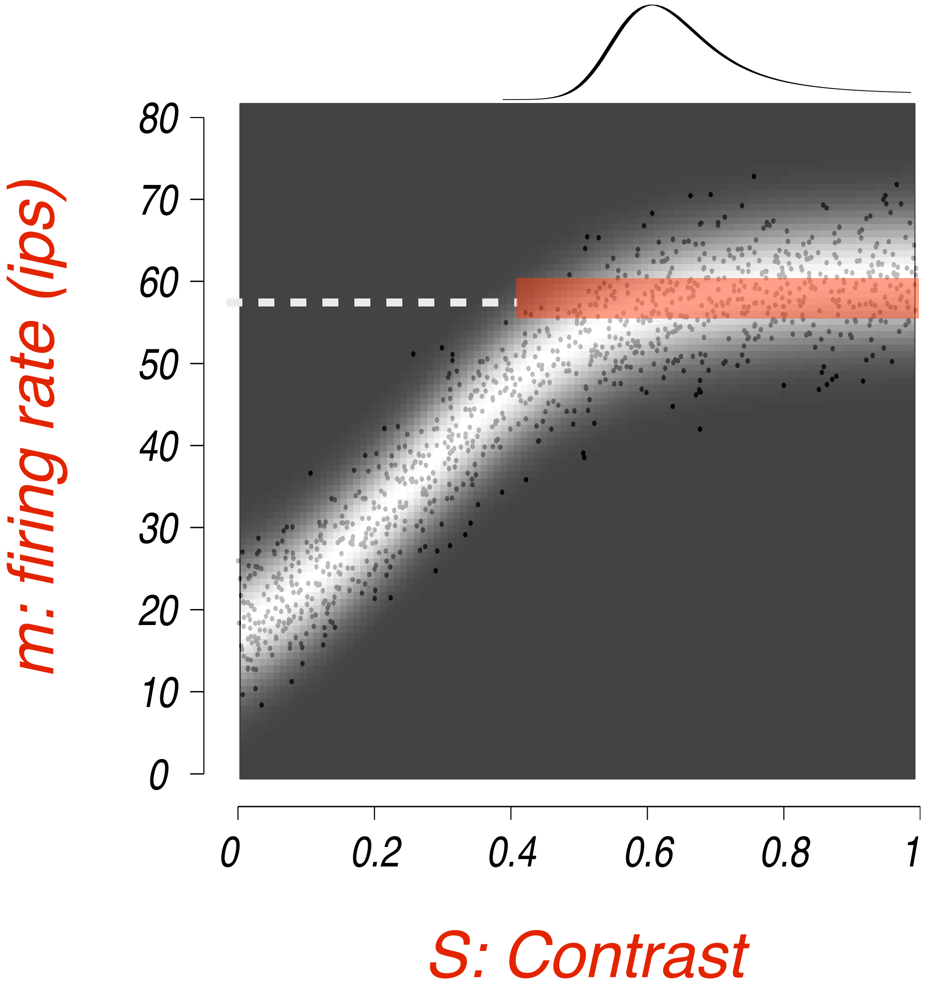
estimator

$$S \xrightarrow{p(m|S)} m^* \xrightarrow{S=f(m)} S^*$$

stimulus

noisy measurement

estimate of the stimulus



Maximum Likelihood Estimation (MLE)

Estimation

noise process

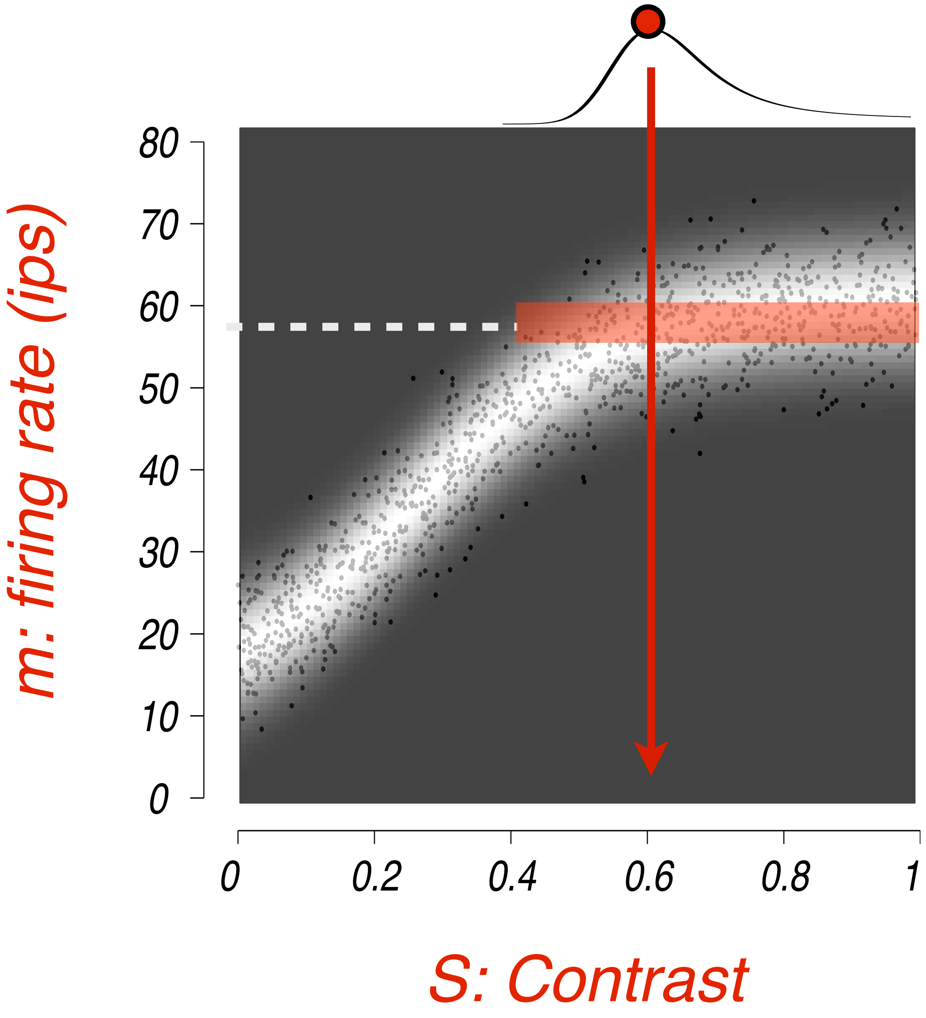
estimator

$$S \xrightarrow{p(m|S)} m^* \xrightarrow{S = f(m)} S^*$$

stimulus

noisy measurement

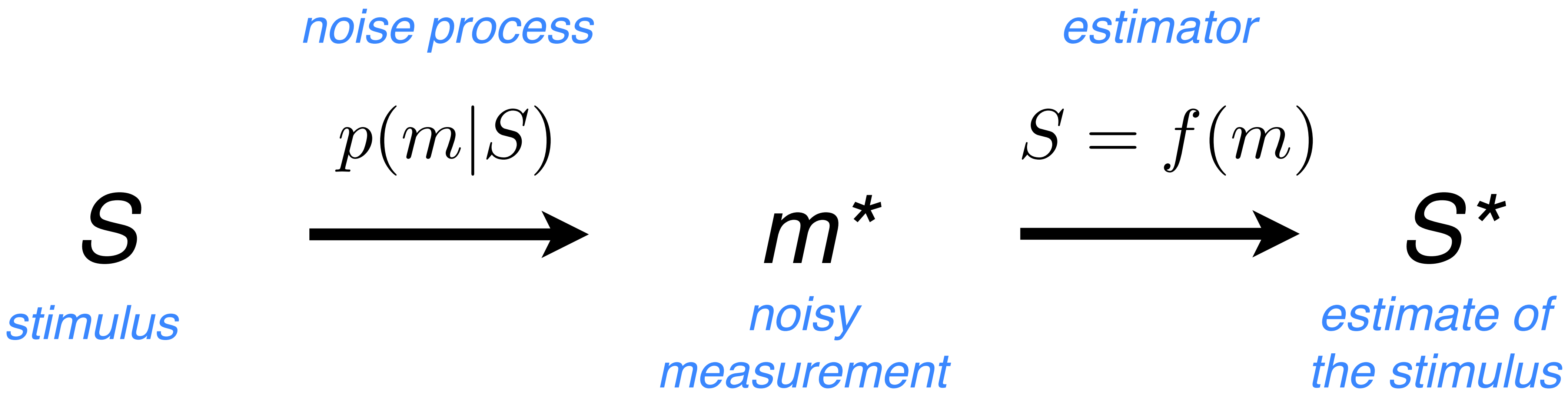
estimate of the stimulus



Maximum Likelihood Estimation (MLE)

$$f_{ML}(m)$$

Example: Plot L(S; 58), and find $f_{ML}(58)$



noise process

$p(m|S)$

estimator

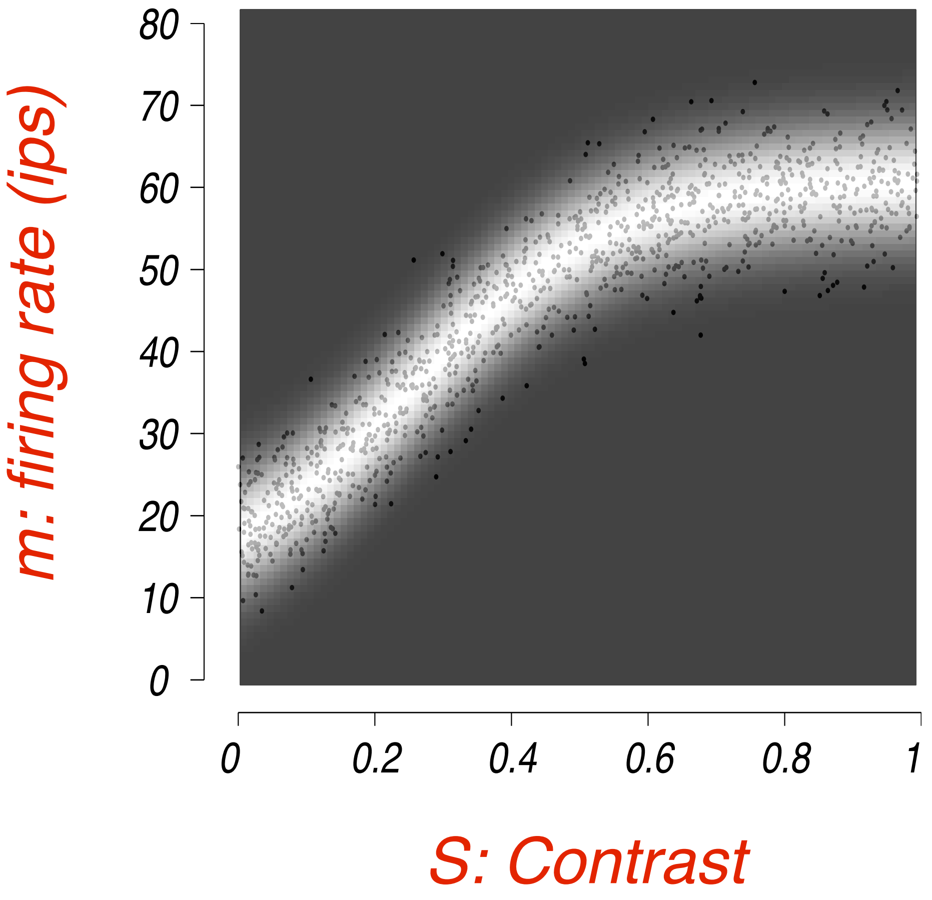
$S = f(m)$

$$S \xrightarrow{p(m|S)} m^* \xrightarrow{S = f(m)} S^*$$

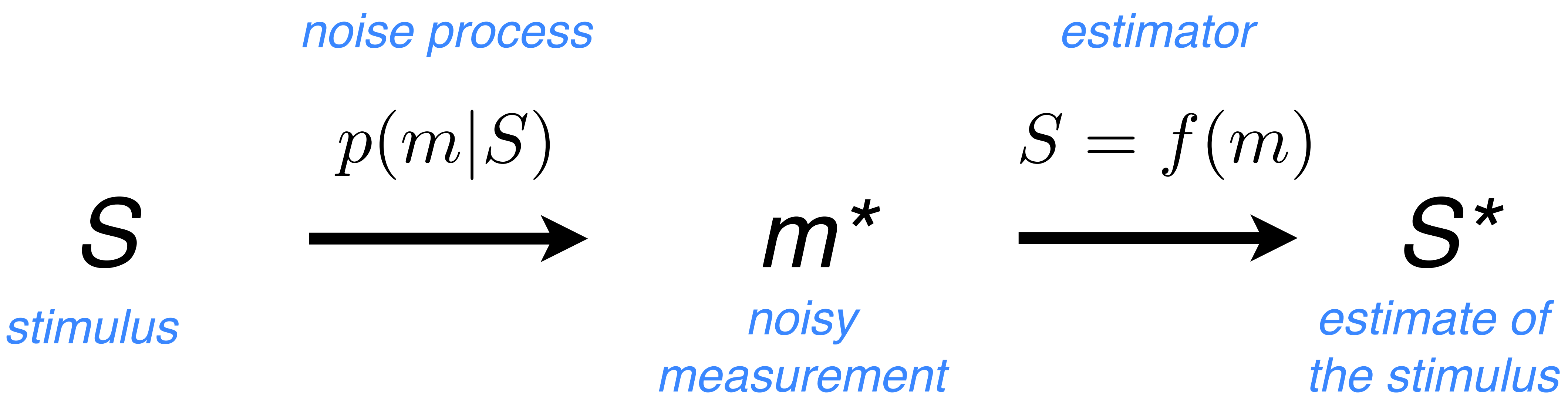
stimulus

noisy measurement

estimate of the stimulus



Example: Plot L(S; 58), and find $f_{ML}(58)$



noise process

$p(m|S)$

estimator

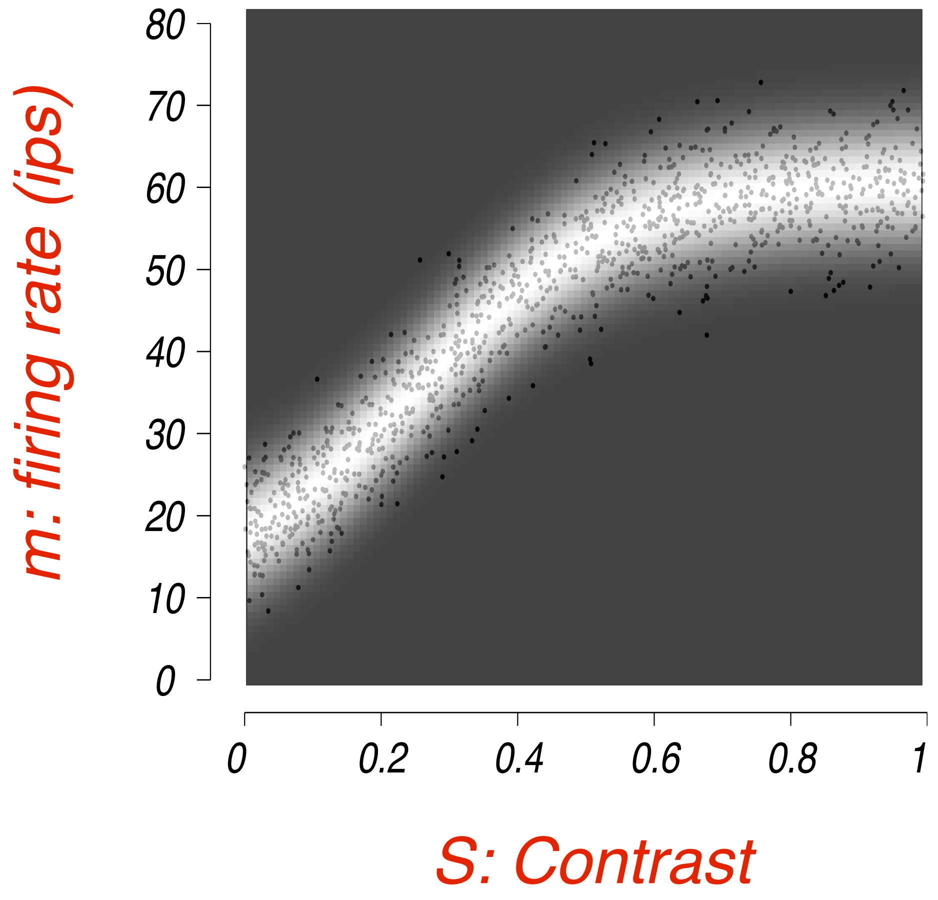
$S = f(m)$

$$S \longrightarrow m^* \longrightarrow S^*$$

stimulus

noisy measurement

estimate of the stimulus



$$m = r(S) + n$$

$$r(S) = 10 + \frac{50}{1 + e^{-\frac{S-0.25}{0.15}}}$$

$$p(n) \approx G(\mu = 0, \sigma = 5)$$

Example: Plot L(S; 58), and find $f_{ML}(58)$

noise process

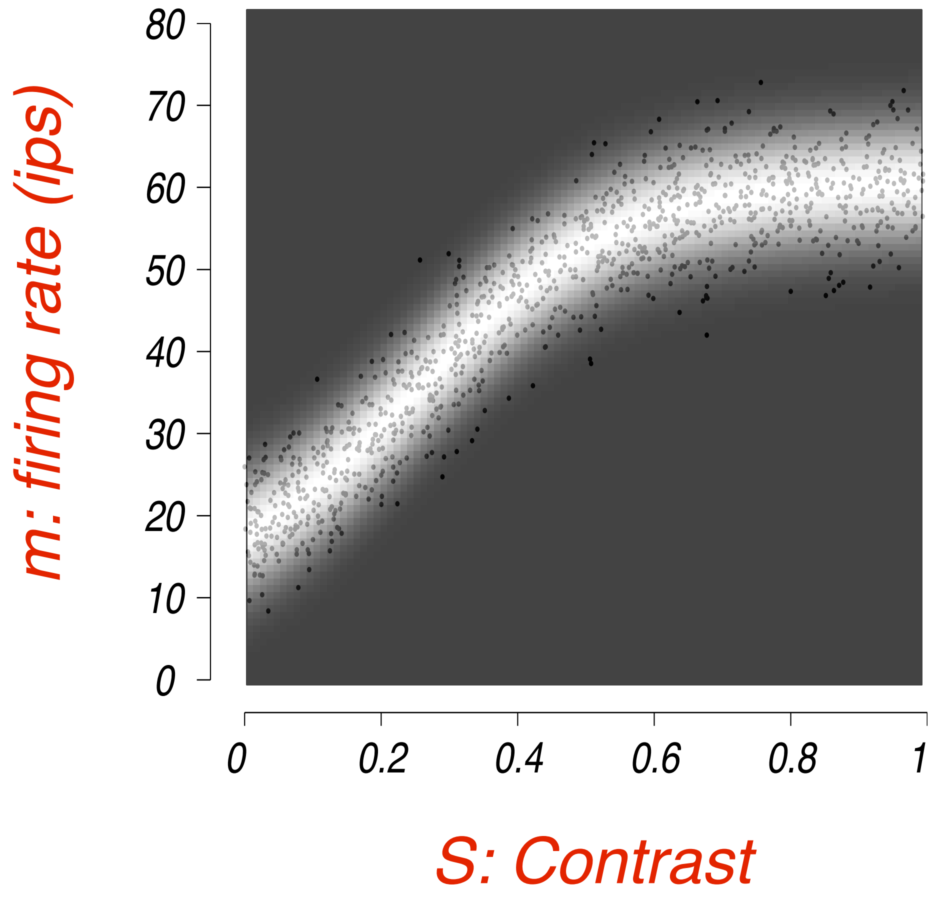
estimator

$$S \xrightarrow{p(m|S)} m^* \xrightarrow{S = f(m)} S^*$$

stimulus

noisy measurement

estimate of the stimulus



$$m = r(S) + n$$

$$r(S) = 10 + \frac{50}{1 + e^{-\frac{S-0.25}{0.15}}}$$

$$p(n) = \frac{1}{\sqrt{2\pi}(5)} e^{-\frac{n^2}{2(5)^2}}$$

Bayesian estimation

noise process

estimator

$$S \xrightarrow{p(m|S)} m^* \xrightarrow{S = f(m)} S^*$$

stimulus

noisy measurement

estimate of the stimulus

A Bayesian estimator is just another f(m). But what it does is that it minimizes some cost over the posterior, p(Slm)

$$p(S|m) = \frac{1}{p(m)} p(m|S) p(S)$$

Posterior

Likelihood

Prior

Bayesian estimation (formal treatment)

Three ingredients for bayesian estimation

1. Likelihood $p(m|S)$
2. Prior $p(S)$

} jointly determine the posterior $p(S|m)$

3. Cost function $C(S_e, S)$ } "cost" of making an estimate *S*

"cost" of making an estimate S_e when the true value is *S*

$$S_e(m) = \arg\min_{S_e} \int C(S_e, S) p(S|m) dS$$

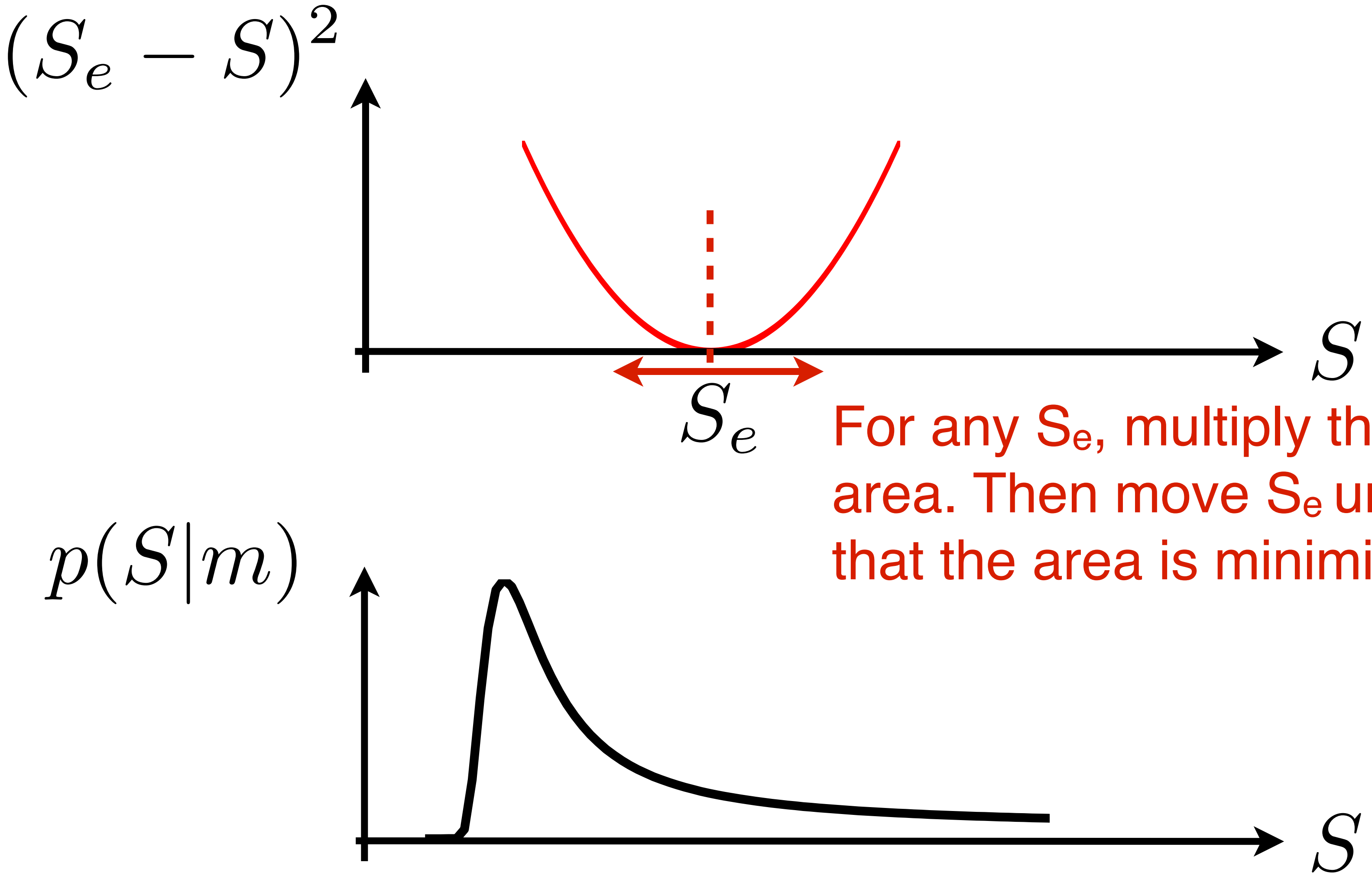
Typical cost functions and Bayesian estimators

squared error cost function

$$C(S_e, S) = (S_e - S)^2$$

need to find S_e that minimizes

$$\int (S_e - S)^2 p(S|m) dS$$



For any S_e, multiply the two curves, find the area. Then move S_e until you find the point that the area is minimized

Typical cost functions and Bayesian estimators

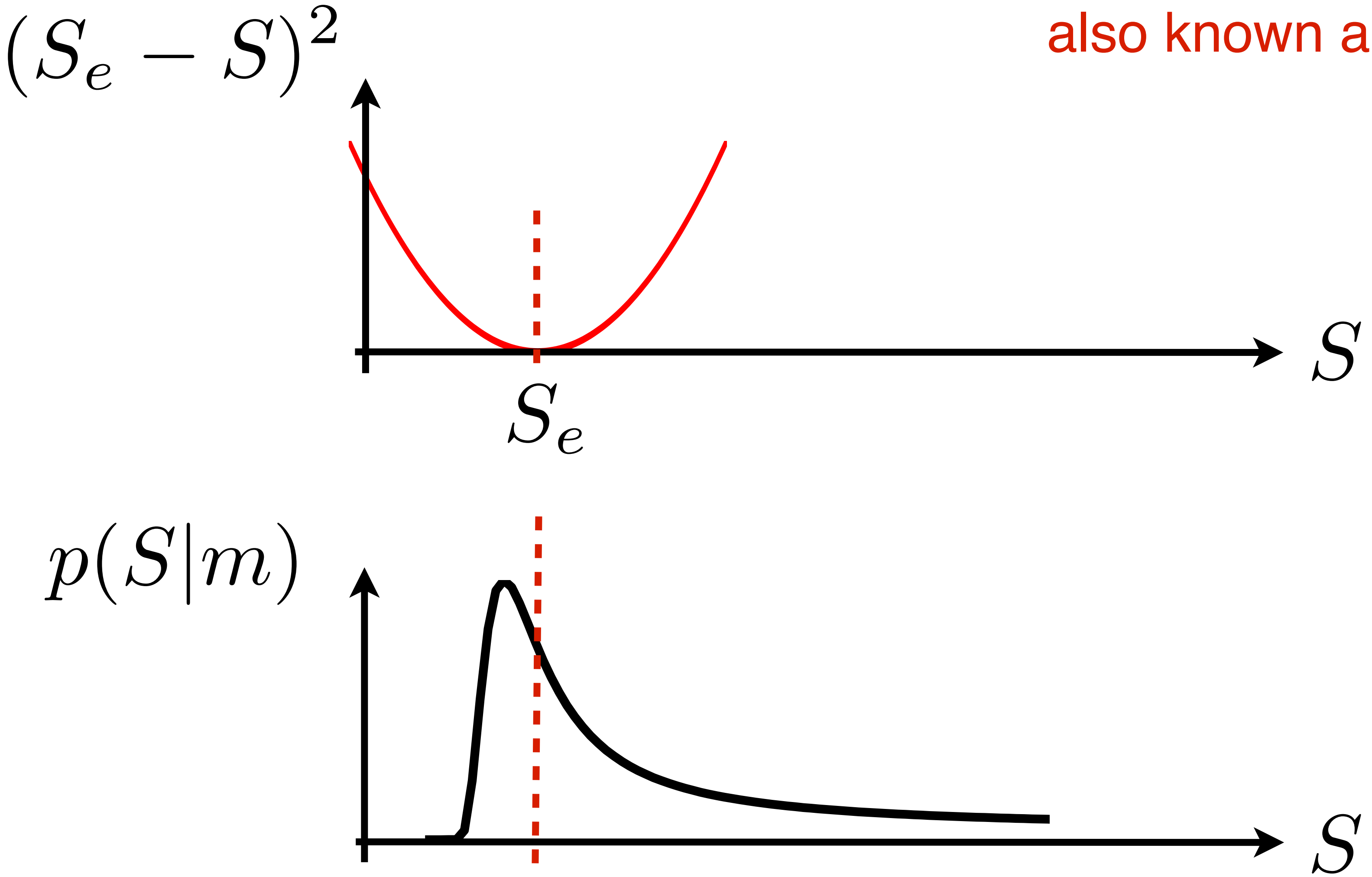
squared error cost function

$$C(S_e, S) = (S_e - S)^2$$

need to find S_e that minimizes

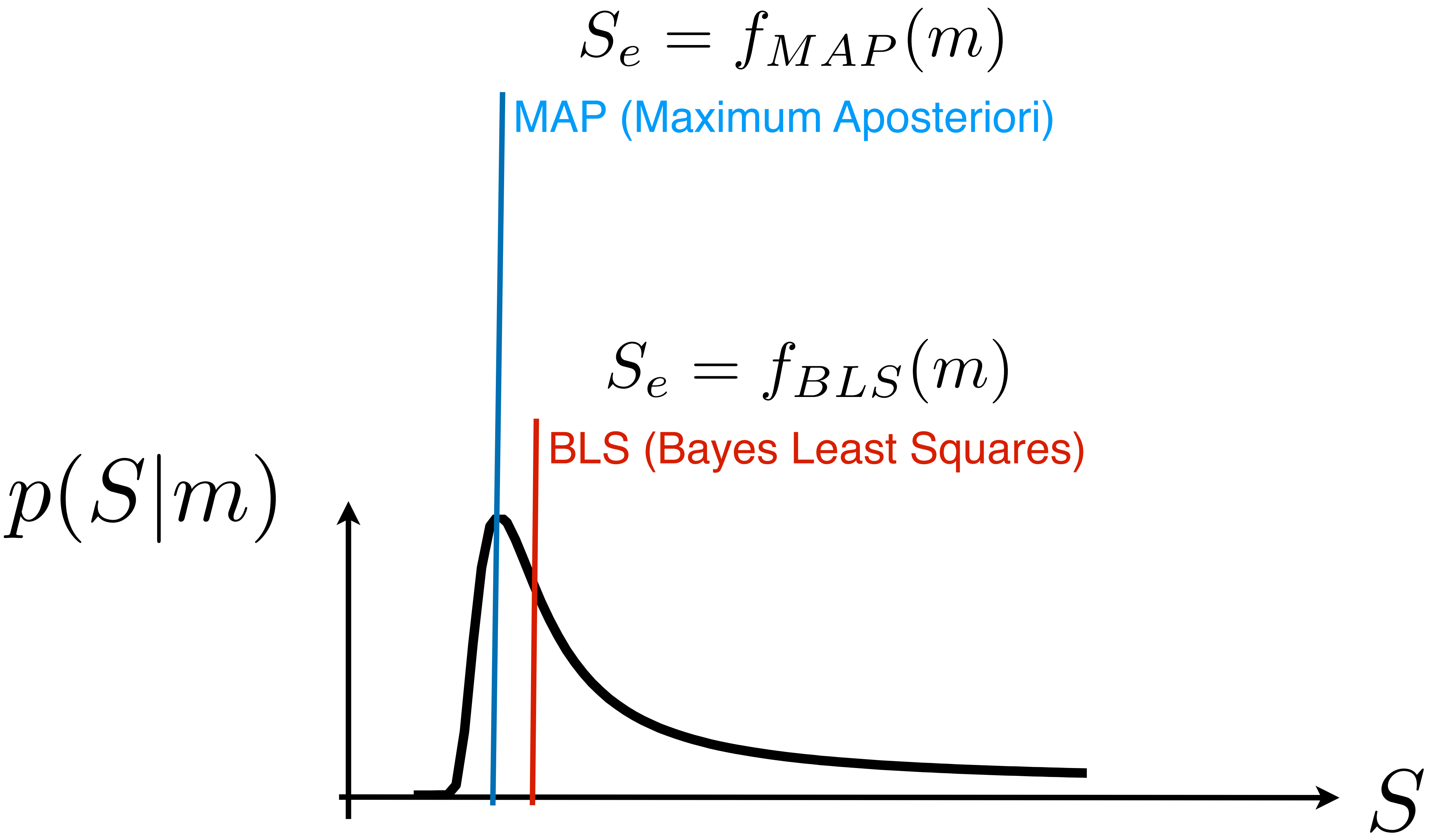
$$\int (S_e - S)^2 p(S|m) dS$$

Bayes Least Squares (BLS)
also known as MMSE



mean of the posterior

Typical cost functions and Bayesian estimators



$$S_e = f_{MAP}(m)$$

MAP (Maximum Aposteriori)

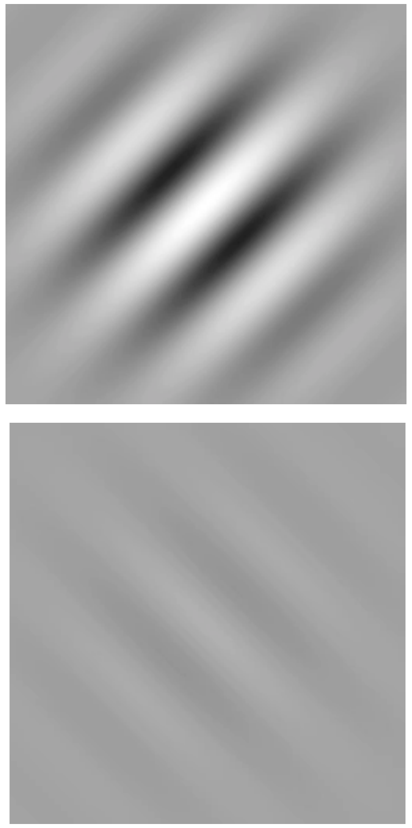
$$S_e = f_{BLS}(m)$$

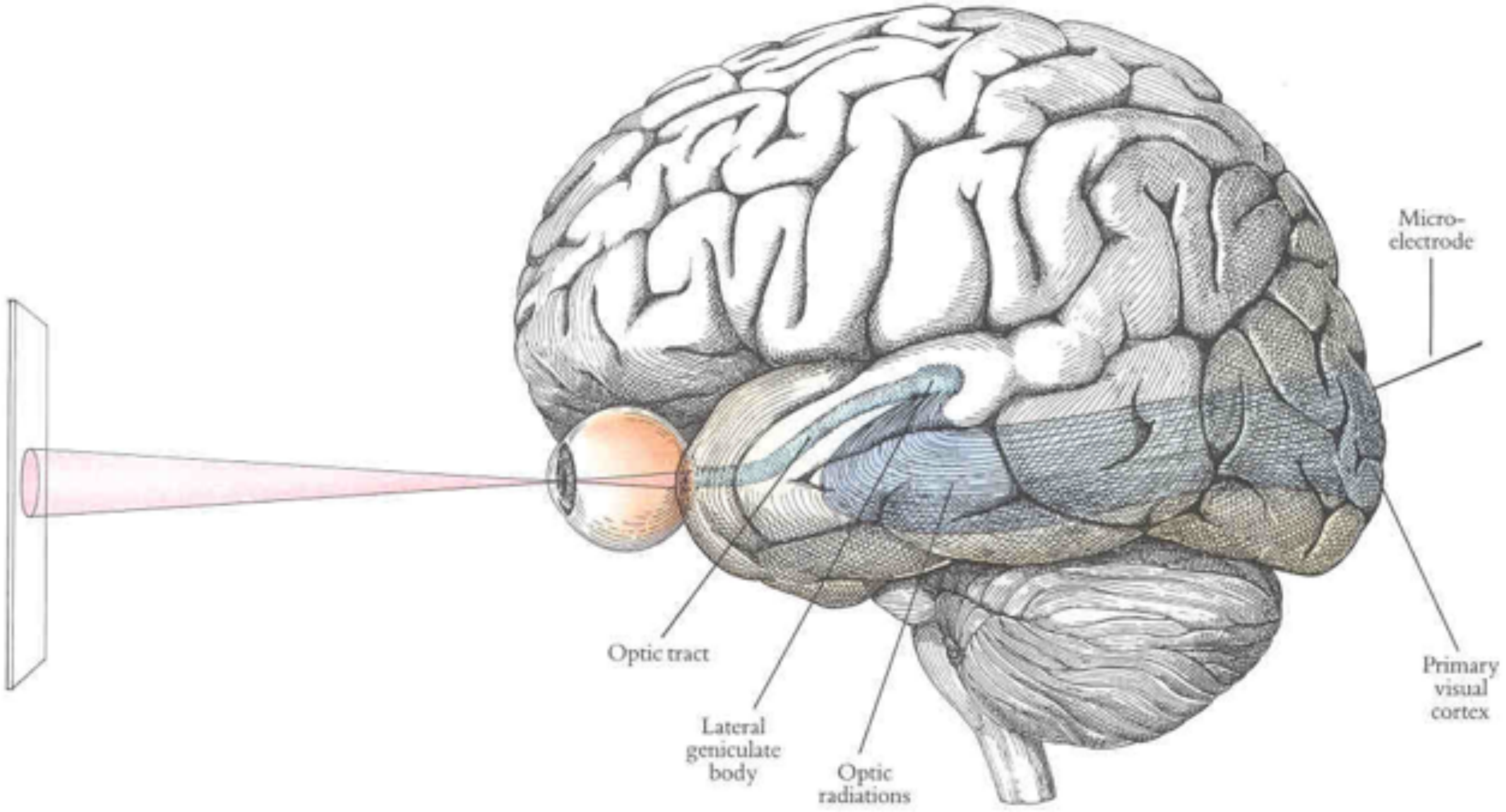
BLS (Bayes Least Squares)

$p(S|m)$

S

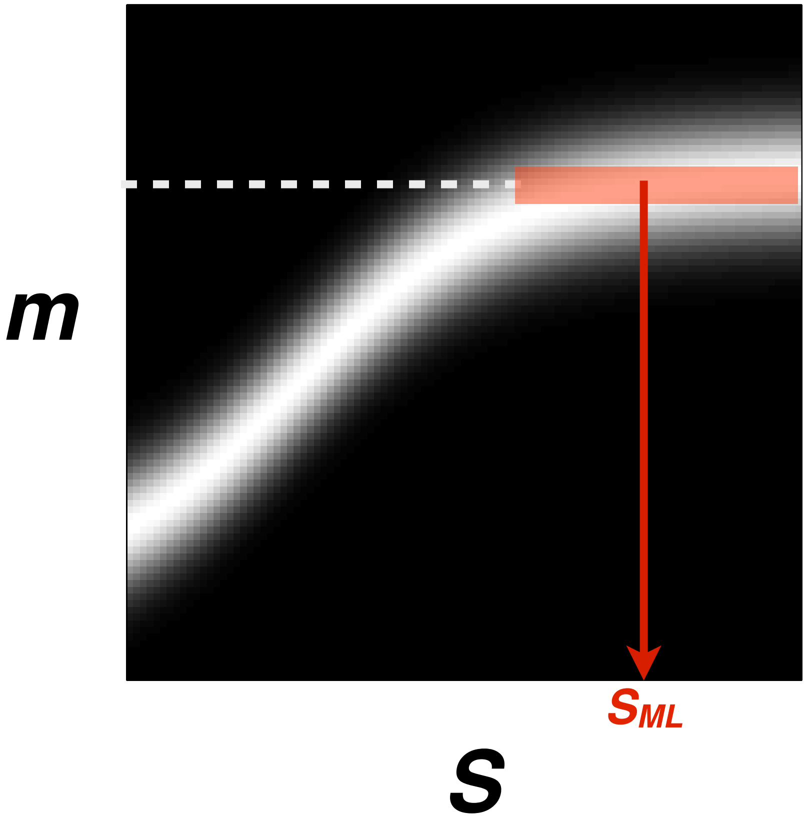
Estimating visual contrast from neural activity





Likelihood

$$p(m|S)$$



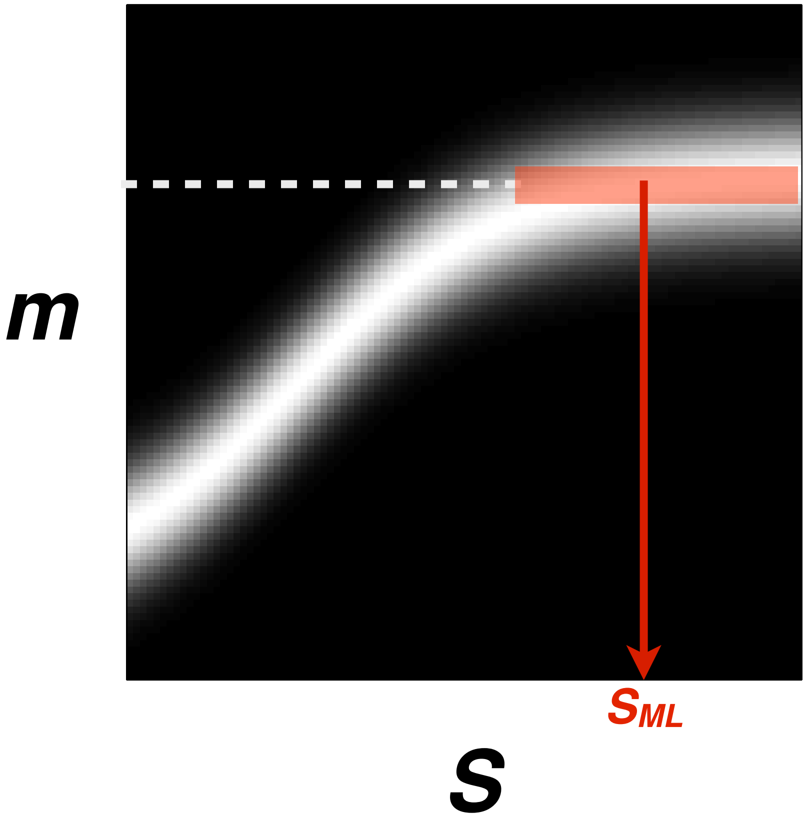
$$m = r(S) + n$$

$$r(S) = 10 + \frac{50}{1 + e^{-\frac{S-0.25}{0.15}}}$$

$$p(n) = \frac{1}{\sqrt{2\pi}(5)} e^{-\frac{n^2}{2(5)^2}}$$

Likelihood

$$p(m|S)$$



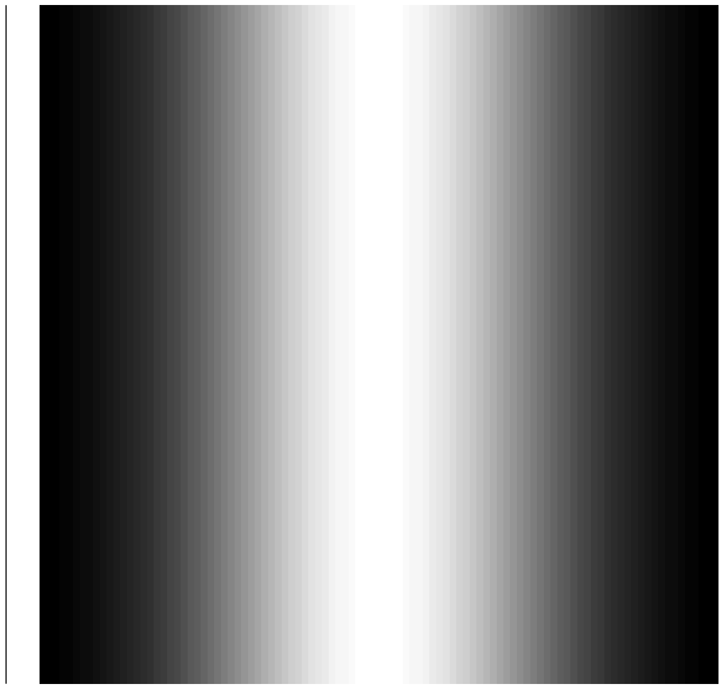
$$m = r(S) + n$$

$$r(S) = 10 + \frac{50}{1 + e^{-\frac{S-0.25}{0.15}}}$$

$$p(n) = \frac{1}{\sqrt{2\pi}(5)} e^{-\frac{n^2}{2(5)^2}}$$

Prior

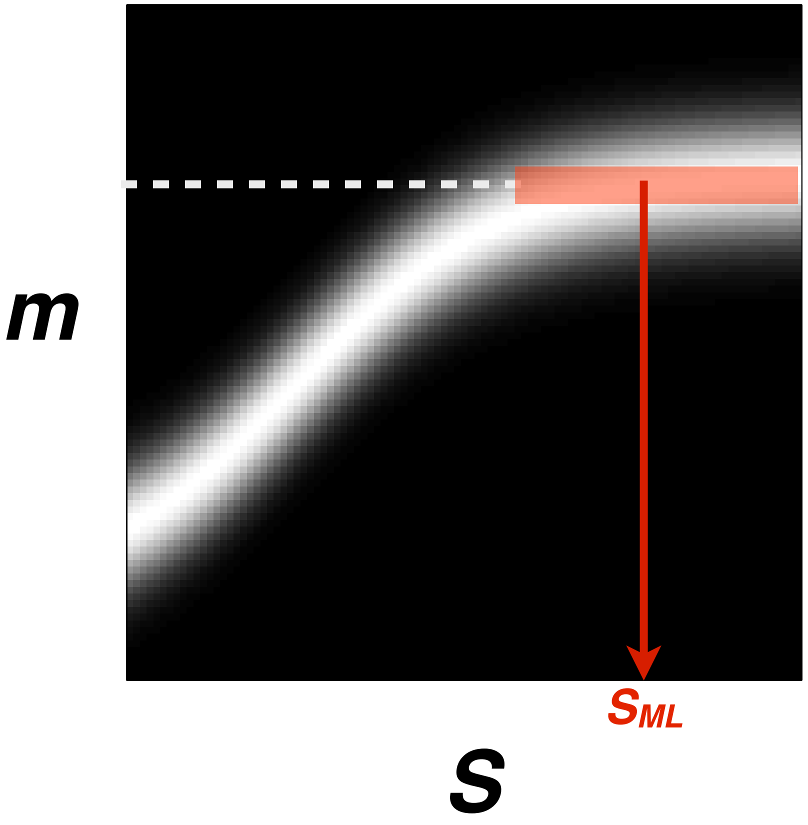
$$\pi(S)$$



$$\pi(S) = \frac{1}{\sqrt{2\pi}(0.2)} e^{-\frac{(S-0.5)^2}{2(0.2)^2}}$$

Likelihood

$$p(m|S)$$



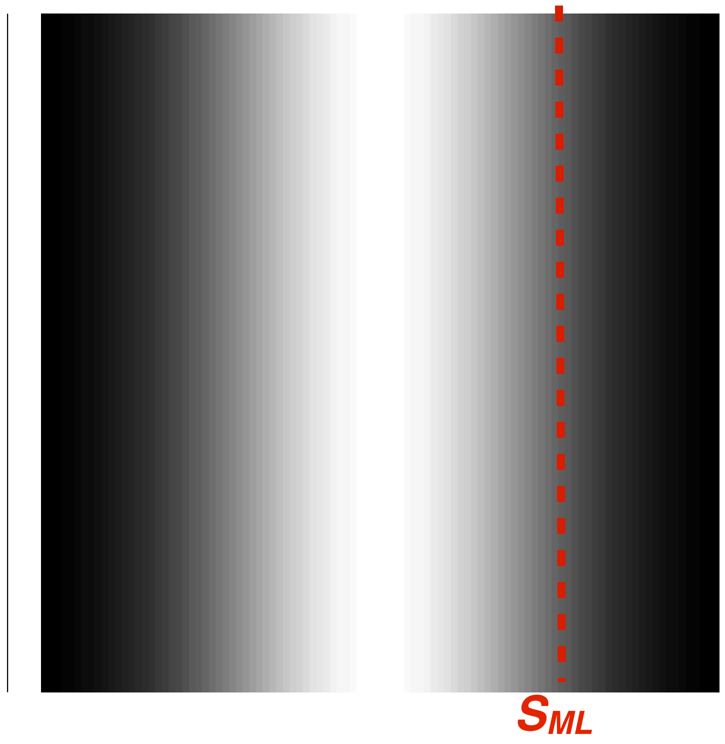
$$m = r(S) + n$$

$$r(S) = 10 + \frac{50}{1 + e^{-\frac{S-0.25}{0.15}}}$$

$$p(n) = \frac{1}{\sqrt{2\pi}(5)} e^{-\frac{n^2}{2(5)^2}}$$

Prior

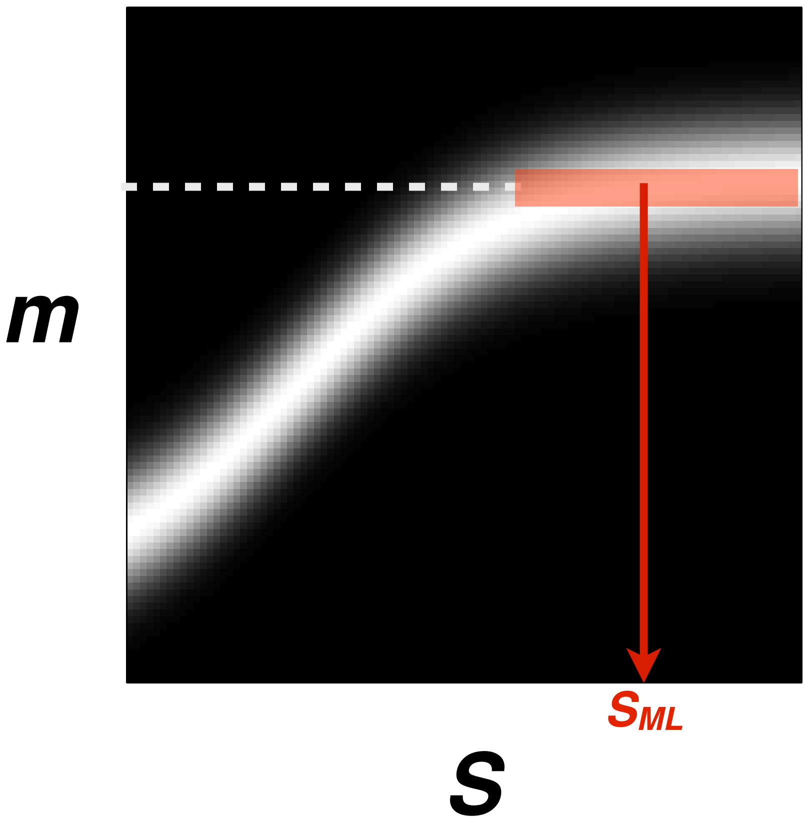
$$\pi(S)$$



$$\pi(S) = \frac{1}{\sqrt{2\pi}(0.2)} e^{-\frac{(S-0.5)^2}{2(0.2)^2}}$$

Likelihood

$$p(m|S)$$



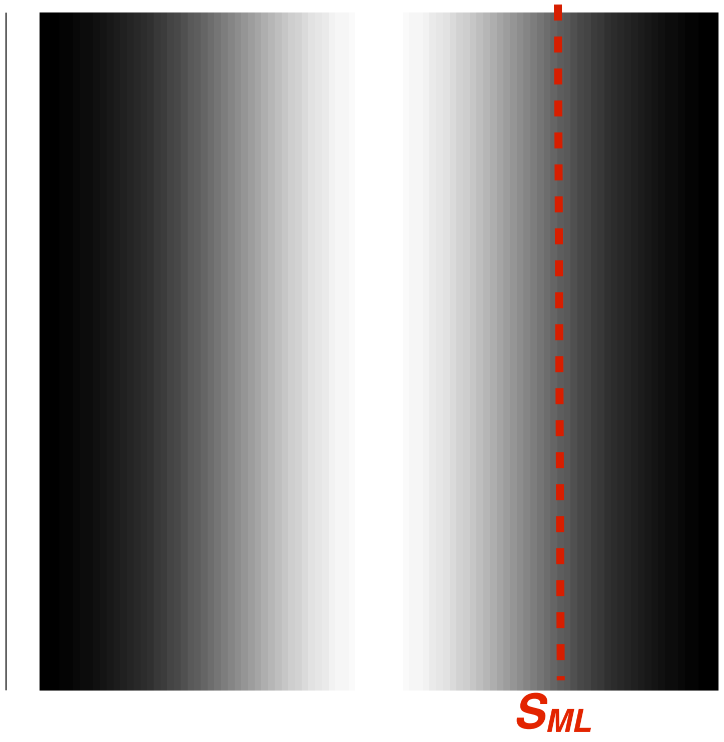
$$m = r(S) + n$$

$$r(S) = 10 + \frac{50}{1 + e^{-\frac{S-0.25}{0.15}}}$$

$$p(n) = \frac{1}{\sqrt{2\pi}(5)} e^{-\frac{n^2}{2(5)^2}}$$

Prior

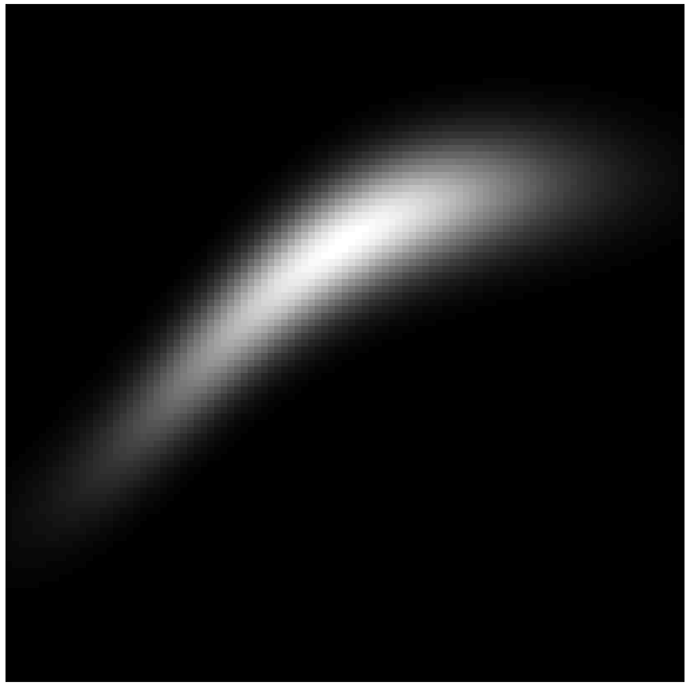
$$\pi(S)$$



$$\pi(S) = \frac{1}{\sqrt{2\pi}(0.2)} e^{-\frac{(S-0.5)^2}{2(0.2)^2}}$$

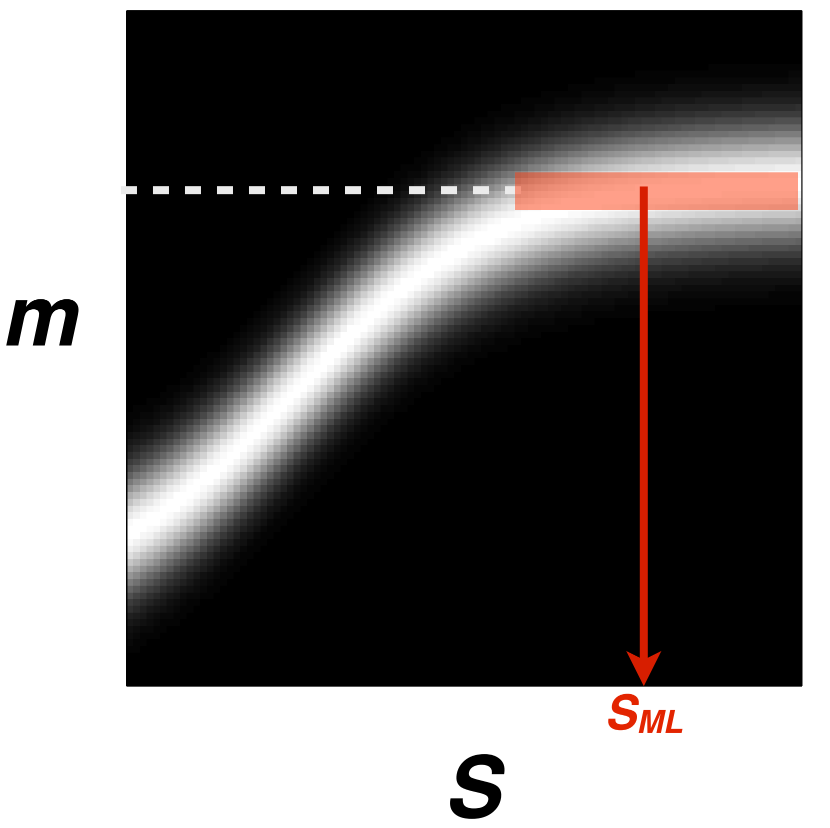
Posterior

$$p(S|m)$$



Likelihood

$$p(m|S)$$



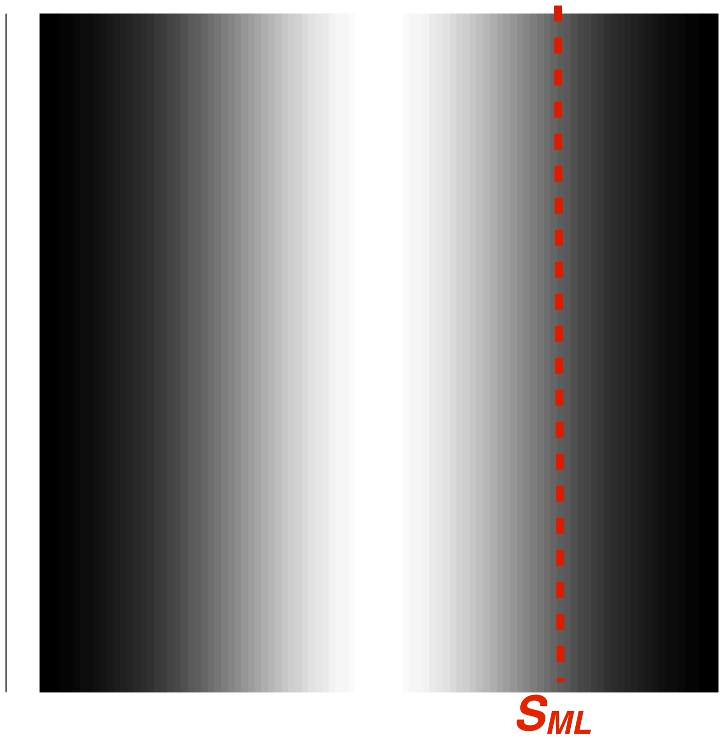
$$m = r(S) + n$$

$$r(S) = 10 + \frac{50}{1 + e^{-\frac{S-0.25}{0.15}}}$$

$$p(n) = \frac{1}{\sqrt{2\pi}(5)} e^{-\frac{n^2}{2(5)^2}}$$

Prior

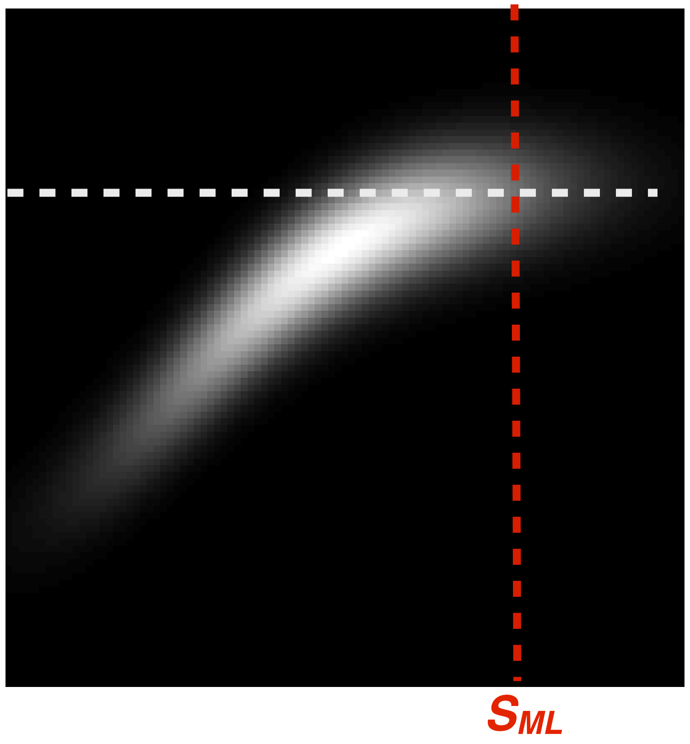
$$\pi(S)$$



$$\pi(S) = \frac{1}{\sqrt{2\pi}(0.2)} e^{-\frac{(S-0.5)^2}{2(0.2)^2}}$$

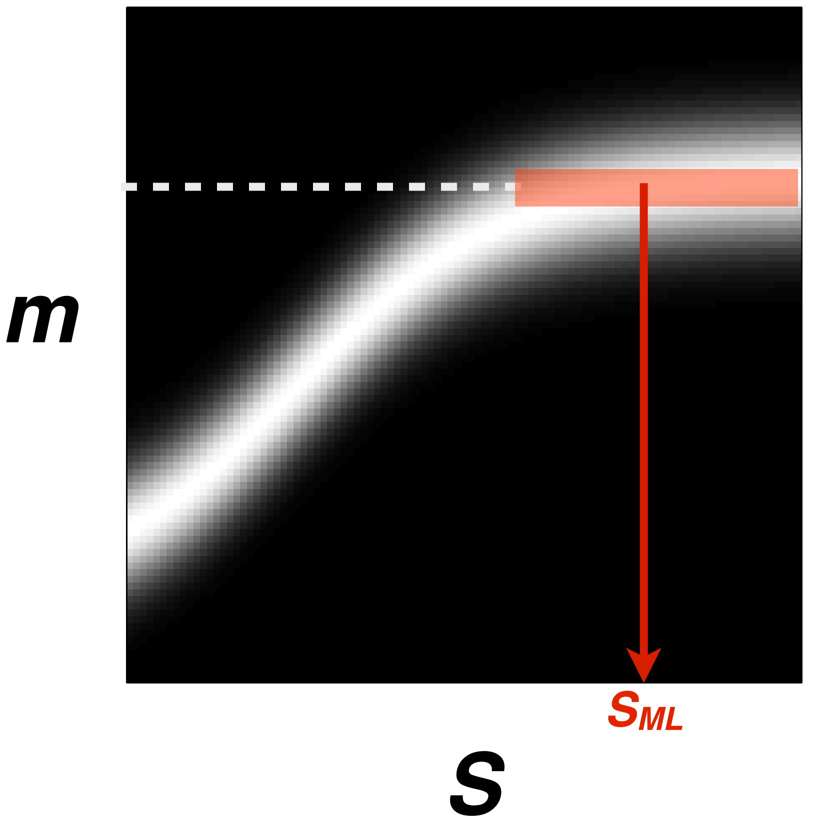
Posterior

$$p(S|m)$$



Likelihood

$$p(m|S)$$



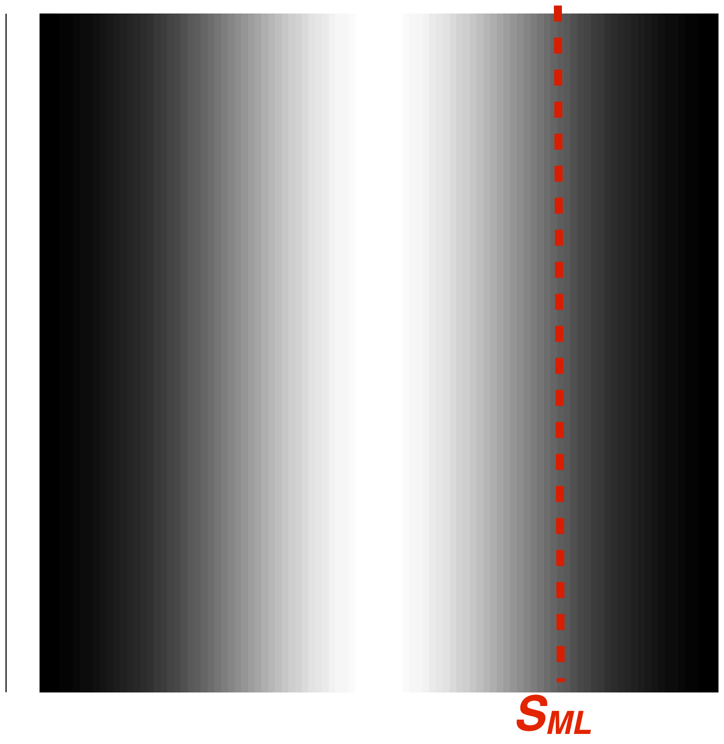
$$m = r(S) + n$$

$$r(S) = 10 + \frac{50}{1 + e^{-\frac{S-0.25}{0.15}}}$$

$$p(n) = \frac{1}{\sqrt{2\pi}(5)} e^{-\frac{n^2}{2(5)^2}}$$

Prior

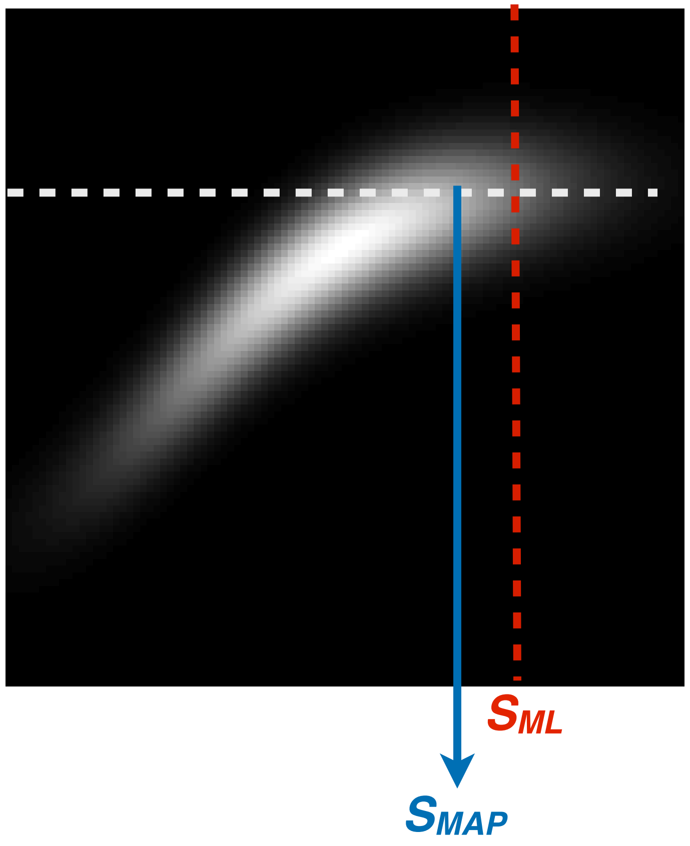
$$\pi(S)$$



$$\pi(S) = \frac{1}{\sqrt{2\pi}(0.2)} e^{-\frac{(S-0.5)^2}{2(0.2)^2}}$$

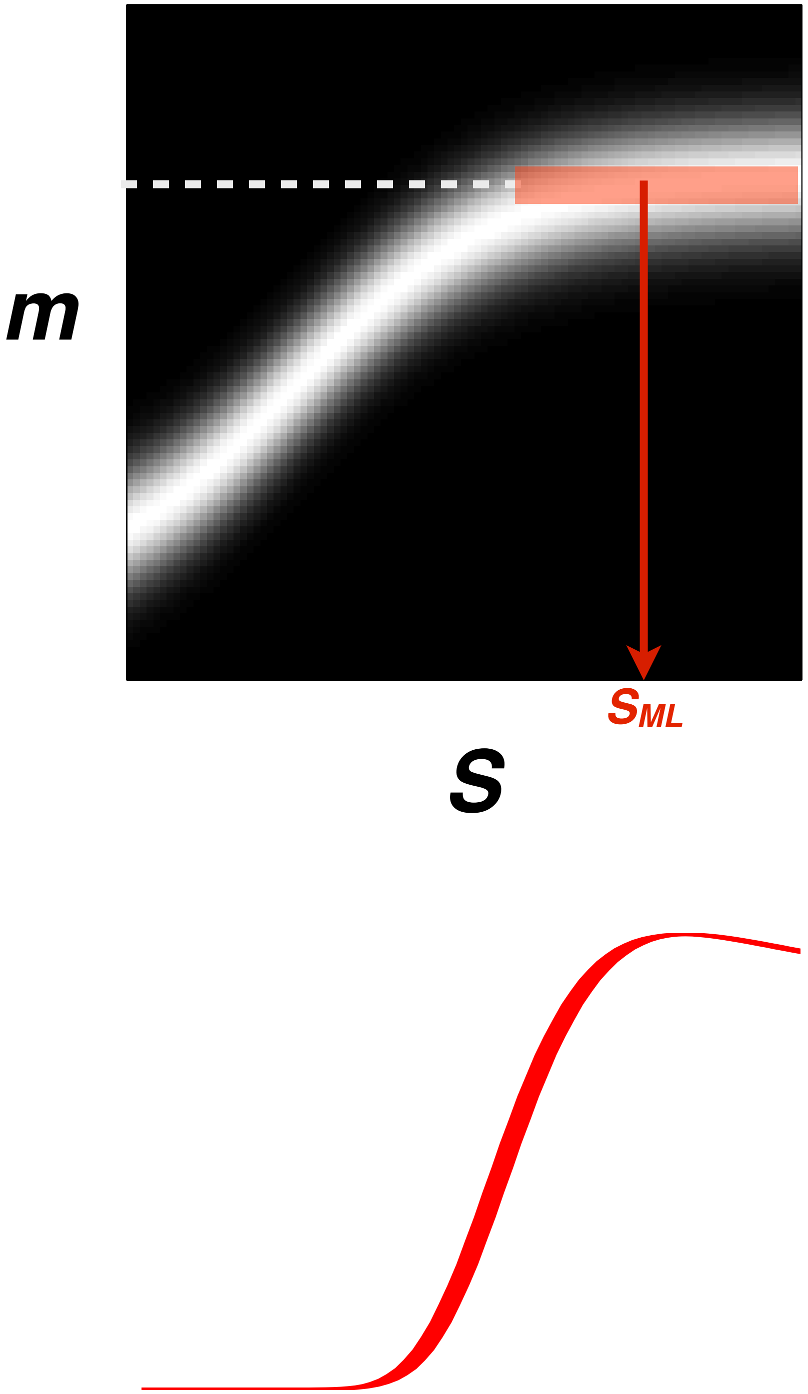
Posterior

$$p(S|m)$$



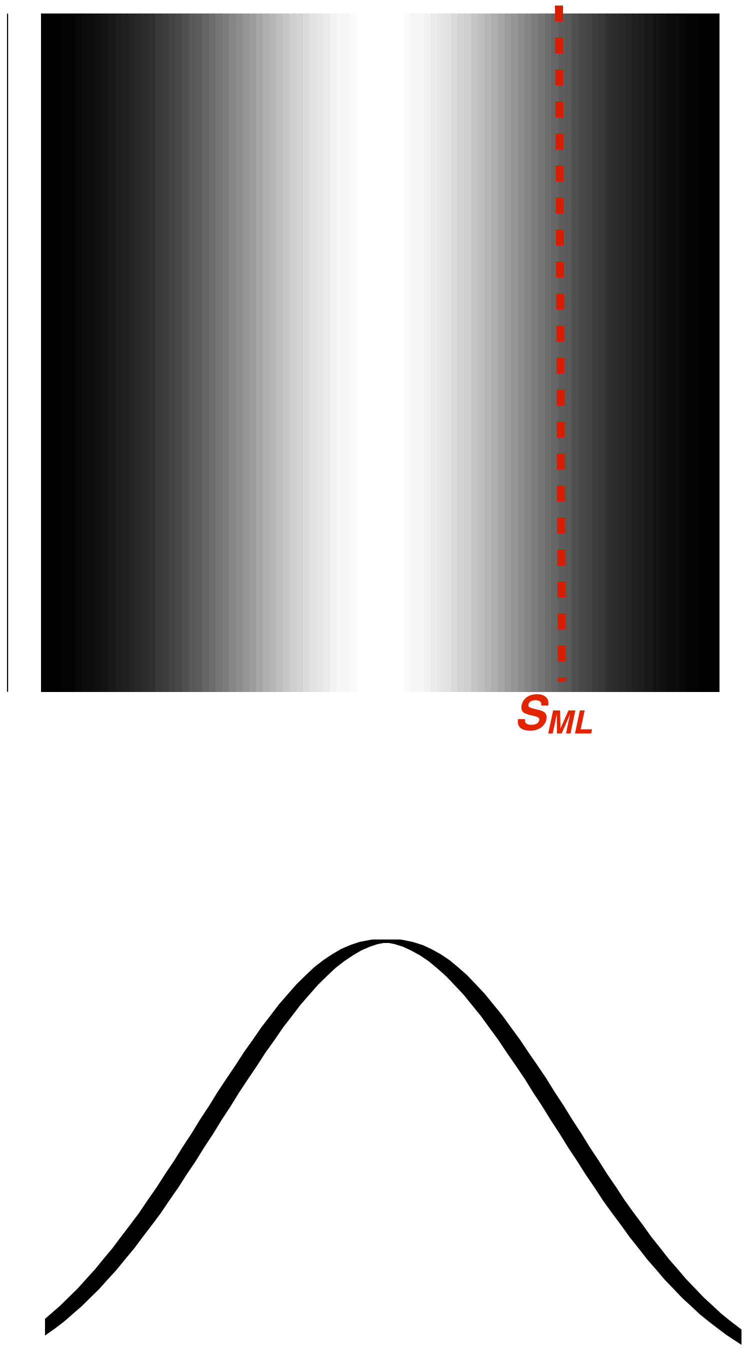
Likelihood

$$p(m|S)$$



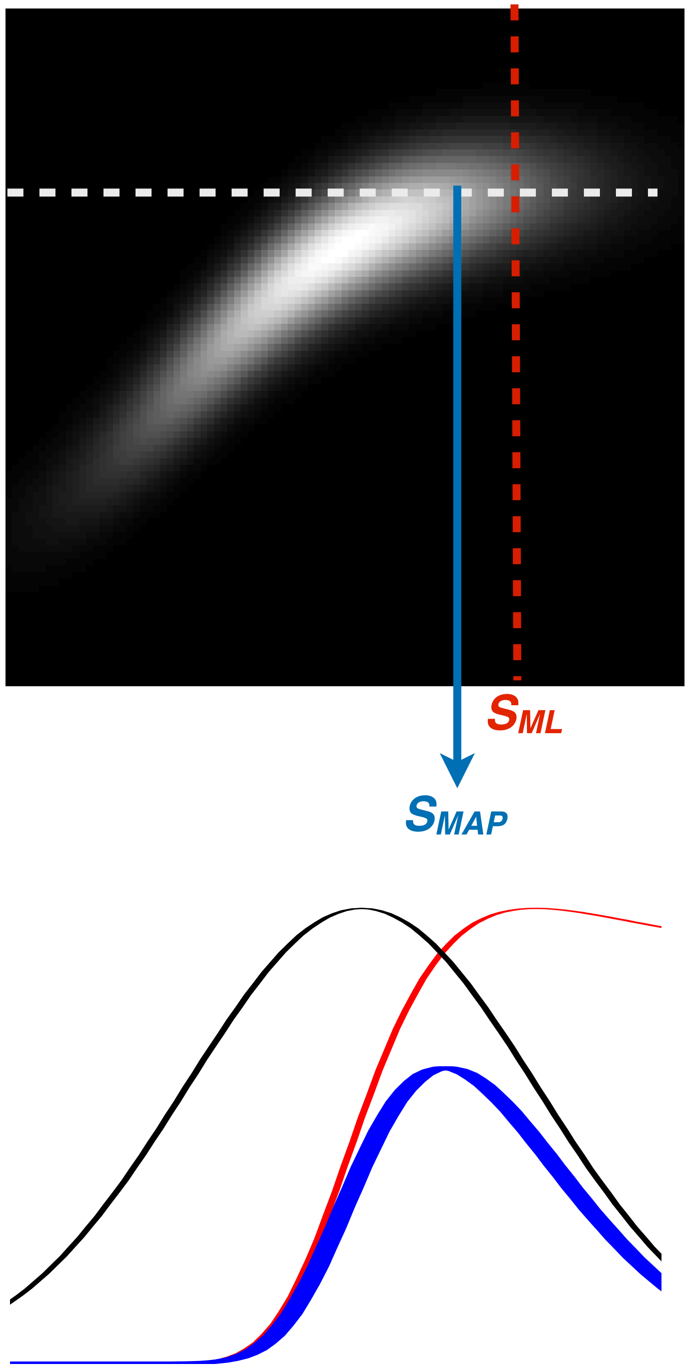
Prior

$$\pi(S)$$



Posterior

$$p(S|m)$$



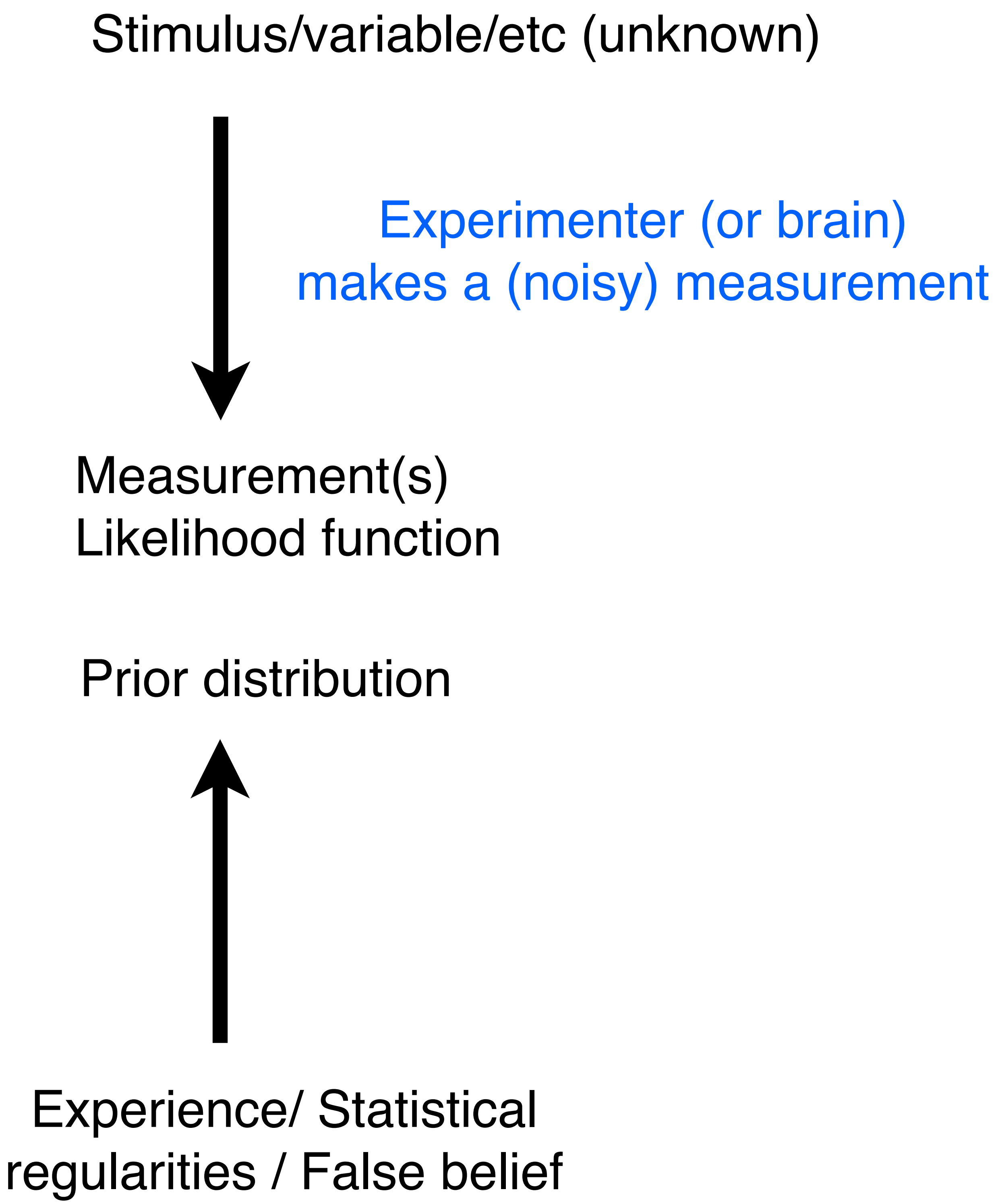
Bayesian estimation

Stimulus/variable/etc (unknown)

Experimenter (or brain)
makes a (noisy) measurement

Measurement(s)
Likelihood function

Bayesian estimation



Stimulus/variable/etc (unknown)

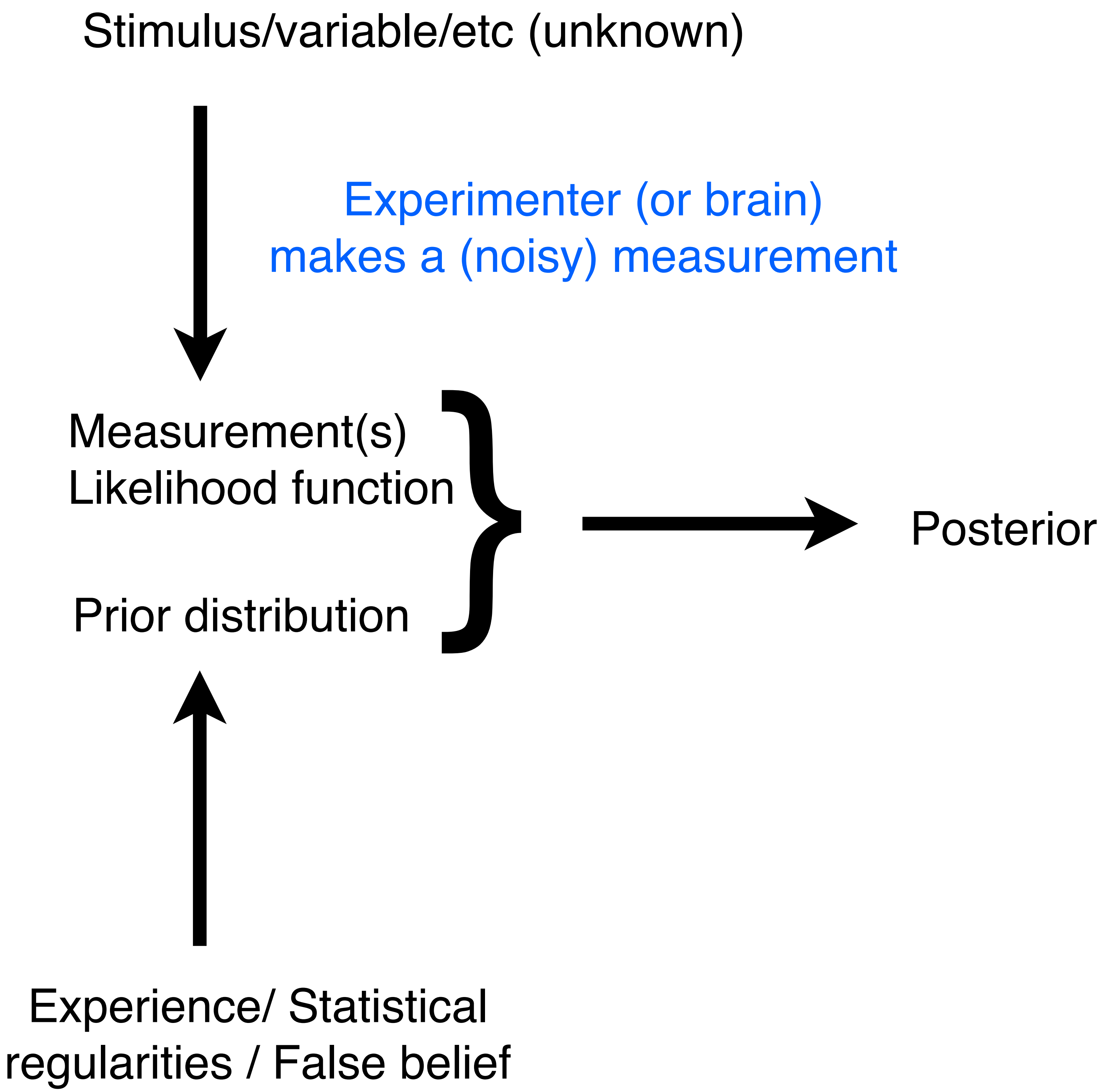
Experimenter (or brain)
makes a (noisy) measurement

Measurement(s)
Likelihood function

Prior distribution

Experience/ Statistical
regularities / False belief

Bayesian estimation



Stimulus/variable/etc (unknown)

Experimenter (or brain)
makes a (noisy) measurement

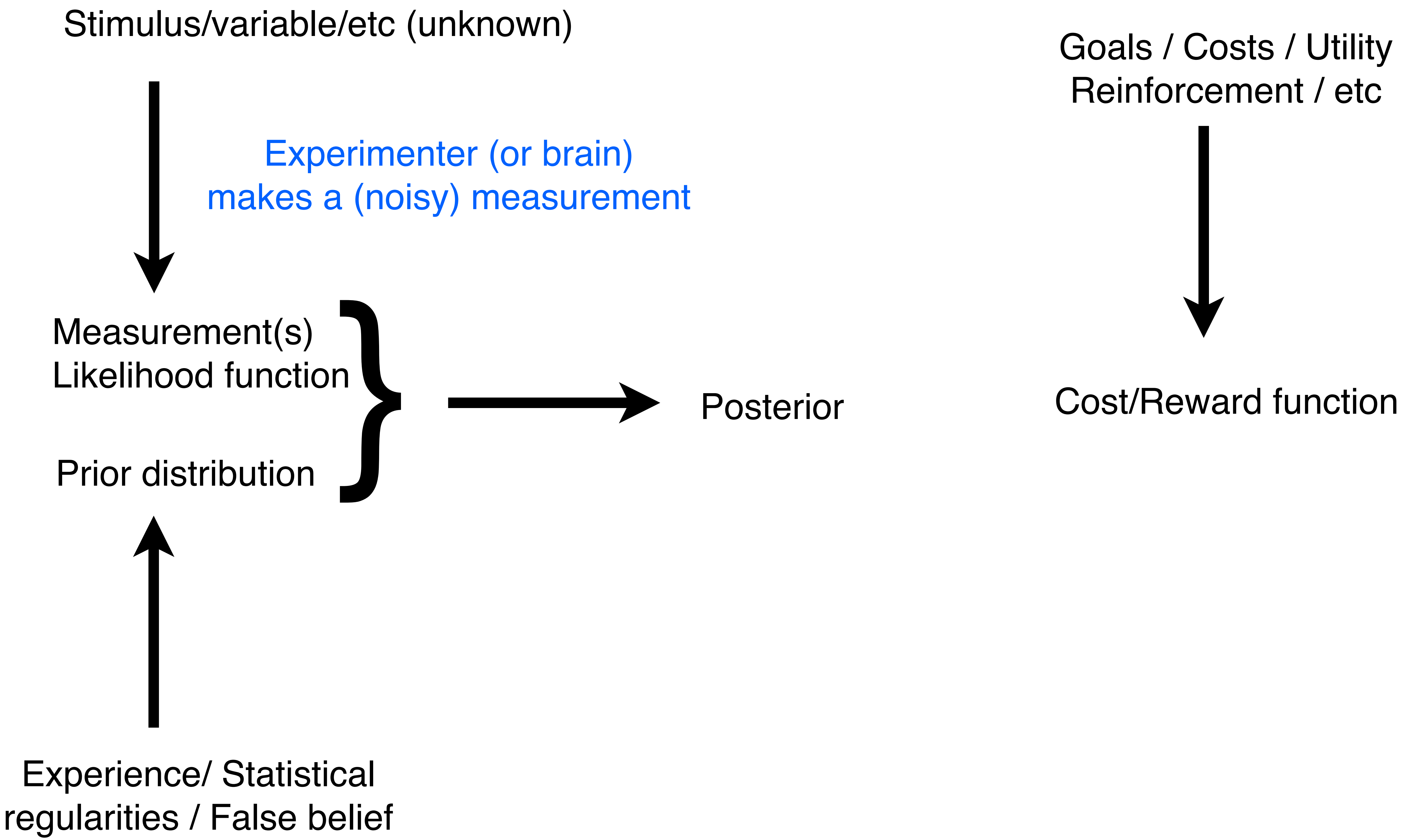
Measurement(s)
Likelihood function

Prior distribution

Posterior

Experience/ Statistical
regularities / False belief

Bayesian estimation



Stimulus/variable/etc (unknown)

Experimenter (or brain)
makes a (noisy) measurement

Measurement(s)
Likelihood function

Prior distribution

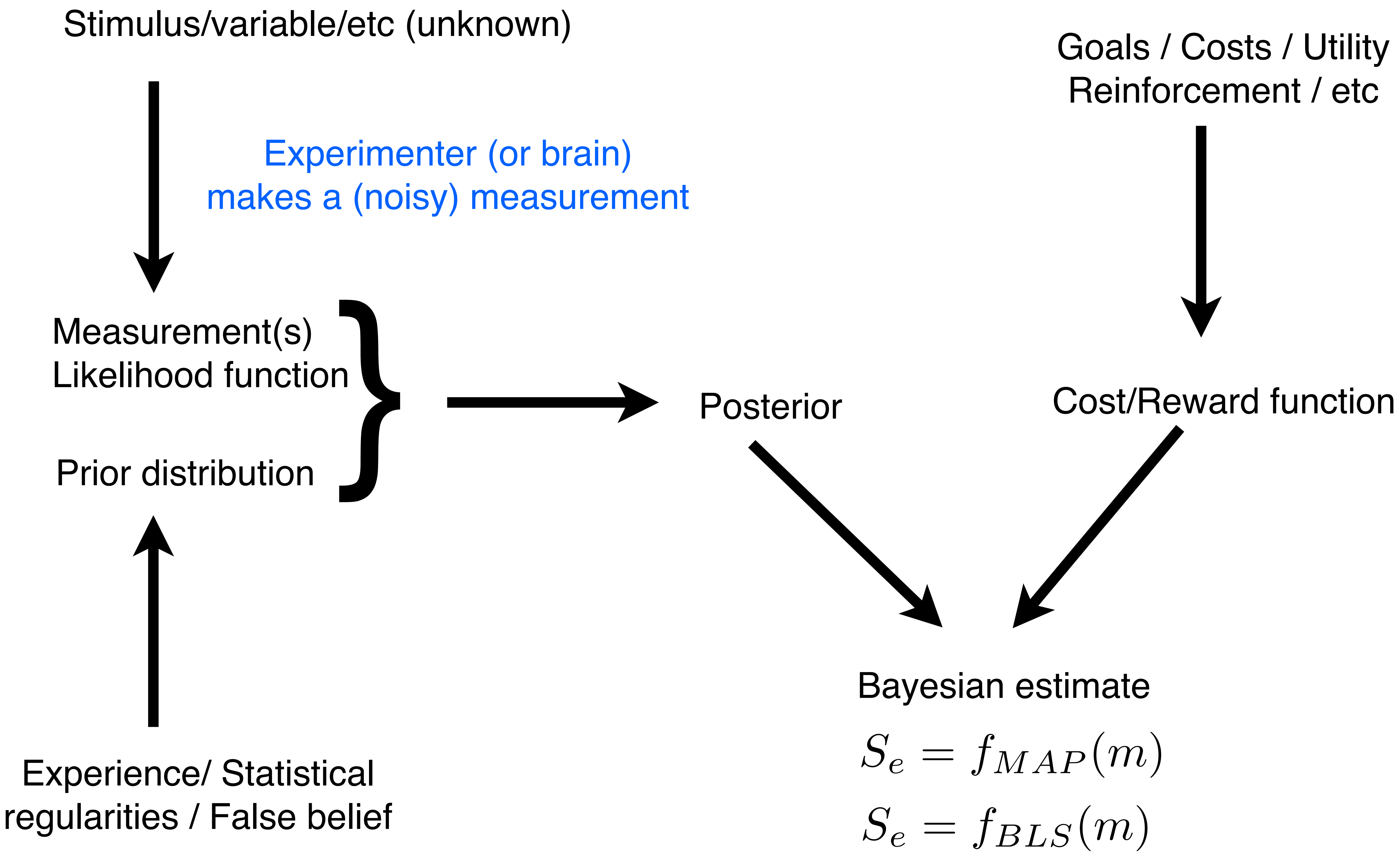
Posterior

Experience/ Statistical
regularities / False belief

Goals / Costs / Utility
Reinforcement / etc

Cost/Reward function

Bayesian estimation



Stimulus/variable/etc (unknown)

Experimenter (or brain)
makes a (noisy) measurement

Measurement(s)
Likelihood function

Prior distribution

Experience/ Statistical
regularities / False belief

Posterior

Goals / Costs / Utility
Reinforcement / etc

Cost/Reward function

Bayesian estimate

$$S_e = f_{MAP}(m)$$

$$S_e = f_{BLS}(m)$$