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

Medical testing
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
\[ p(Yes) \quad p(+|Yes) \quad p(-|No) \]
\[
p(Yes) \quad p(+|Yes) \quad p(-|No)
\]

\[2\% \quad 99\% \quad 90\%\]

rare
| $p(\text{Yes})$ | $p(+|\text{Yes})$ | $p(-|\text{No})$ |
|----------------|--------------------|------------------|
| 2%            | 99%                | 90%              |

| 1000 | 20 | $\sim 20^+/20^+$ | $882^-/980^-$   |
|      |    |                  | $98^+/980^-$    |

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

\[
p(Yes|+) = \frac{20}{20+98} = 0.17
\]

17%
\[ p(Yes) \quad p(\,+|Yes) \quad p(-|No) \]

\[
2\% \quad 99\% \quad 90\%
\]

\[
p(Yes|+) = \frac{p(\,+|Yes)p(Yes)}{p(+)}
\]

\[
99\% \quad 2\% \\
10\% \quad 98\%
\]

\[
p(+) = p(\,+|Yes)p(Yes) + p(\,+|No)p(No)
\]
Estimating visual contrast from neural activity

Which grating moves faster?

Application #1: Biases in Motion Perception
Estimating visual contrast from neural activity

\[ S \xrightarrow{\text{noise process}} m^* \xrightarrow{\text{estimator}} S^* \]

- **S**: stimulus
- **m^***: noisy measurement
- **S^***: estimate of the stimulus

Application #1: Biases in Motion Perception
Conditional probability & likelihood function

\[ p(m|S) \]

S: Contrast

m*: noisy measurement

S: stimulus

noise process
Conditional probability & likelihood function

\[ p(m|S) \]

\[ S \rightarrow m^* \]

stimulus

noise process

noisy measurement

S: Contrast

m: firing rate (ips)
Conditional probability & likelihood function

$S$: Contrast

$m$: firing rate (ips)

noise process

$p(m|S)$

$m^*$

noisy measurement

Conditional probability

$p(m|S)$
Conditional probability & likelihood function

\[ p(m|S) \]

\( S \)  
stimulus

\( m^* \)  
noisy measurement

\( S: \) Contrast

\( m: \) firing rate (ips)
Conditional probability & likelihood function

\[ p(m|S) \]

\[ S \xrightarrow{\text{noise process}} m^* \]

\( S: \text{Contrast} \)

\( m: \text{firing rate (ips)} \)

Thursday, June 11, 15
Conditional probability & likelihood function

\[ p(m|S) \]

- \( S \): Contrast
- \( m \): firing rate (ips)

Noise process

\( m^* \)

Noisy measurement

Conditional probability

\[ p(m|S) \]
Conditional probability & likelihood function

\[ p(m|S) \]

\[ L(S; m) \]

S: Contrast

m*: noisy measurement

stimulus

noise process

\[ \text{S: Contrast} \]

\[ \text{m: firing rate (ips)} \]
Conditional probability & likelihood function

\[ L(S; m) = p(m|S) \]

\[ p(m|S) \]

\[ S \rightarrow m^* \]

\[ m: \text{firing rate (ips)} \]

\[ S: \text{Contrast} \]

Noise process

stimulus

noisy measurement
Conditional probability & likelihood function

\[ p(m|S) \]

**stimulus**

\[ S \]

**noise process**

\[ m^* \]

**noisy measurement**

Likelihood

\[ L(S; m) = p(m|S) \]
Conditional probability & likelihood function

\[ \text{noise process} \]

\[ p(m|S) \]

\[ m^* \]

\[ \text{stimulus} \]

\[ S \]

\[ \text{noisy measurement} \]

\[ L(S; m) = p(m|S) \]

\[ S: \text{Contrast} \]

\[ m: \text{firing rate (ips)} \]
Estimation

\[ S \rightarrow m^* \rightarrow S^* \]

\[ p(m|S) \]

\[ S = f(m) \]

S: Contrast

m: firing rate (ips)

noise process

estimator

stimulus

noisy measurement

estimate of the stimulus
Estimation

\[ S \xrightarrow{\text{noise process}} m^* \xrightarrow{\text{estimator}} S^* \]

- \( S \): Contrast
- \( m \): firing rate (ips)

Maximum Likelihood Estimation (MLE)
Estimation

\[ S \xrightarrow{\text{noise process}} m^* \rightarrow S^* \]

- \( S \): Contrast
- \( m \): firing rate (ips)

Maximum Likelihood Estimation (MLE)

\[ f_{ML}(m) \]
Example: Plot \( L(S; 58) \), and find \( f_{ML}(58) \)

\[ S \xrightarrow{\text{noise process}} p(m|S) \xrightarrow{\text{estimator}} m^* \xrightarrow{\text{noisy measurement}} S^* \]

\( S \): Contrast

\( m \): firing rate (ips)
Example: Plot $L(S; 58)$, and find $f_{ML}(58)$

\[ S \xrightarrow{\text{noise process}} m^* \xrightarrow{\text{estimator}} S^* \]

- $S$: Contrast
- $m$: firing rate (ips)

\[
m = r(S) + n
\]

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

\[
p(n) \approx G(\mu = 0, \sigma = 5)
\]
Example: Plot $L(S; 58)$, and find $f_{ML}(58)$

\[
m = r(S) + n
\]

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

\[
p(n) = \frac{1}{\sqrt{2\pi(5)}} e^{-\frac{n^2}{2(5)^2}}
\]
A Bayesian estimator is just another \( f(m) \). But what it does is that it minimizes some cost over the posterior, \( p(S|m) \)

\[
p(S|m) = \frac{1}{p(m)} p(m|S)p(S)
\]
Bayesian estimation (formal treatment)

Three ingredients for bayesian estimation

1. Likelihood \[ p(m|S) \]
2. Prior \[ p(S) \]
3. Cost function \[ C(S_e, S) \]

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

jointly determine the posterior \( p(S|m) \)

“cost” of making an estimate \( S_e \) when the true value is \( S \)
Typical cost functions and Bayesian estimators

squadrred 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

\( S \)

\( S_e \)

\( p(S|m) \)

Mean of the posterior

\( S \)
Typical cost functions and Bayesian estimators

\[ S_e = f_{MAP}(m) \]

MAP (Maximum Aposteriori)

\[ S_e = f_{BLS}(m) \]

BLS (Bayes Least Squares)
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

\[ p(m|S) \]

Prior

\[ \pi(S) \]

\[ m = r(S) + n \]

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

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

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

\[ p(m|S) \]

Prior

\[ \pi(S) \]

Posterior

\[ p(S|m) \]

\[ m = r(S) + n \]

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

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

\[ \pi(S) = \frac{1}{\sqrt{2\pi(0.2)}} e^{-\frac{(S-0.5)^2}{2(0.2)^2}} \]
$p(m|S) = r(S) + n$

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

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

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

\[ p(S|m) \]

\[ p(S|m) = \frac{1}{\sqrt{2\pi}(5)} e^{-\frac{(S-0.5)^2}{2(0.2)^2}} \]
**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

Experience/ Statistical regularities / False belief

Posterior
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 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

\[ S_e = f_{\text{MAP}}(m) \]

\[ S_e = f_{\text{BLS}}(m) \]