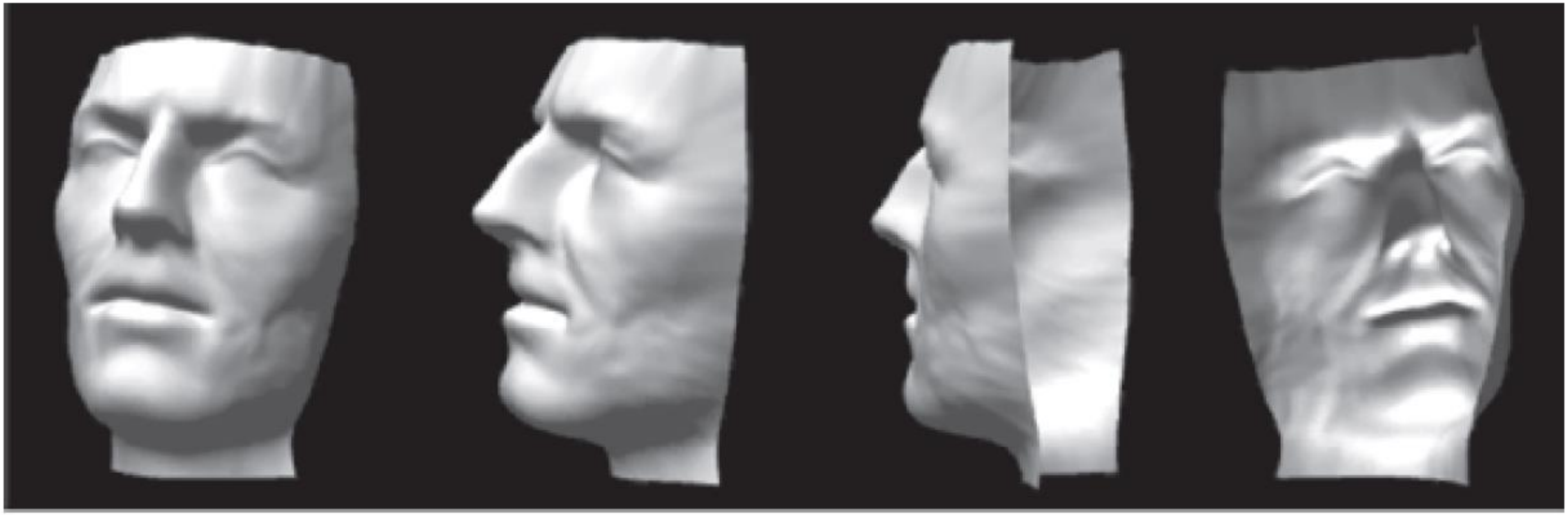
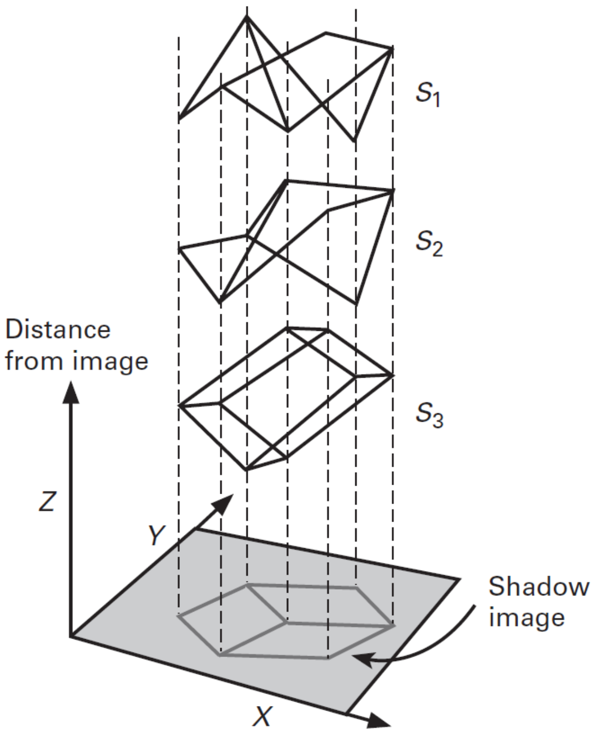
Bayesian Learning



Seeing is not a direct apprehension of reality, as we often like to pretend. Quite contrary, *seeing is inference from incomplete information*…

E. T. Jaynes, 2003.

Vision and Brain by J. V. Stone, 2012



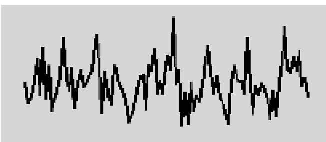
Experience is required to disambiguate images: previous experience is used to interpret ambiguous retinal images by *Bayesian analysis*, for example.

Vision and Brain by J. V. Stone, 2012

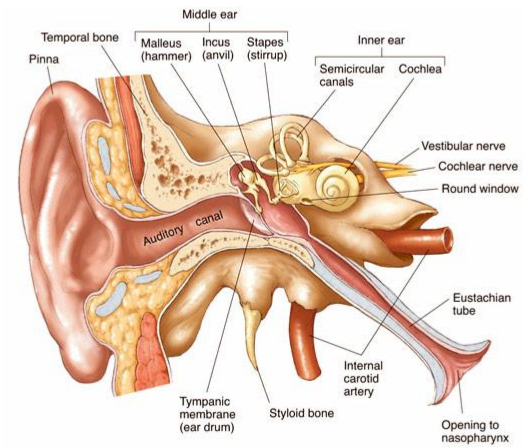
Bayes' rule is a rigorous method for interpreting evidence in the context of previous experience or knowledge.

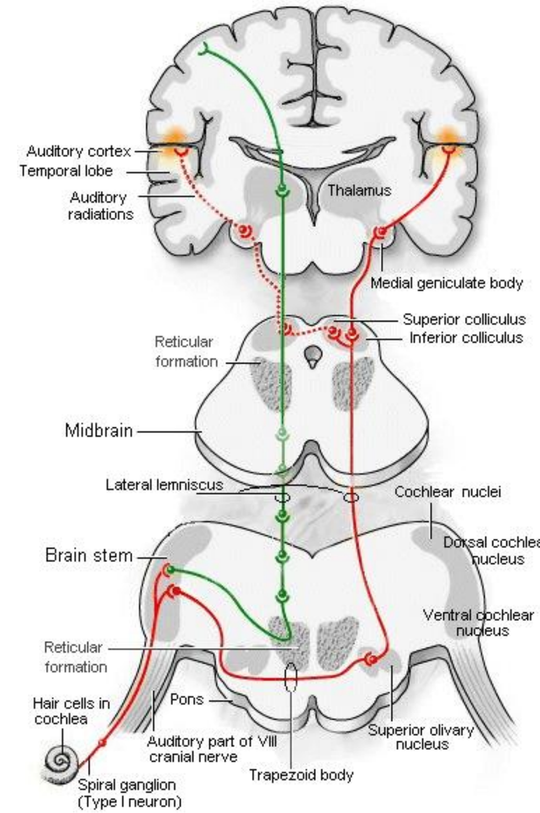
Thomas Bayes (1701 – 1761)   Pierre-Simon Laplace (1749 – 1827)

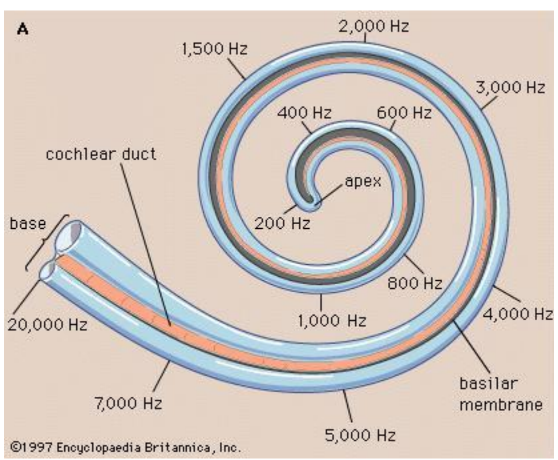
Hearing

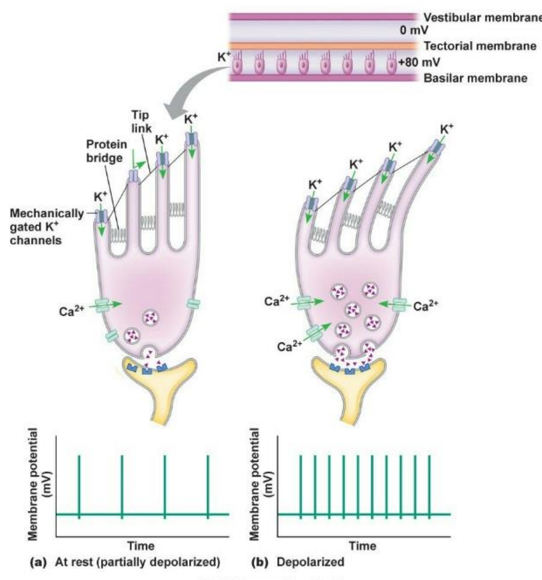


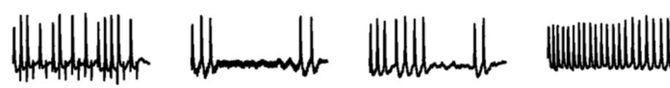
Acoustic Signal





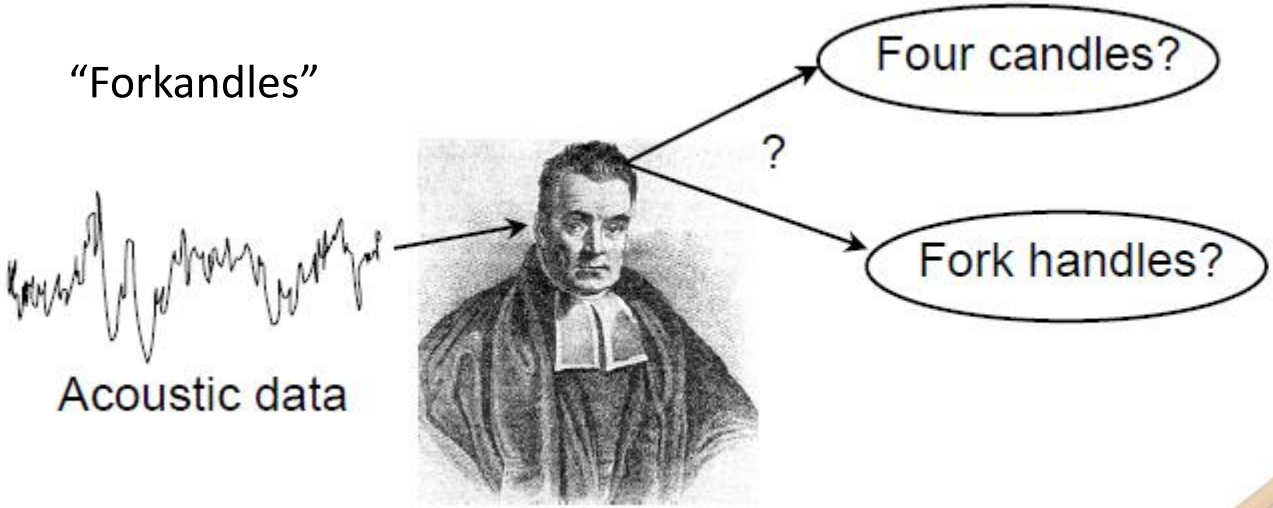






Hearing “Forkandles”

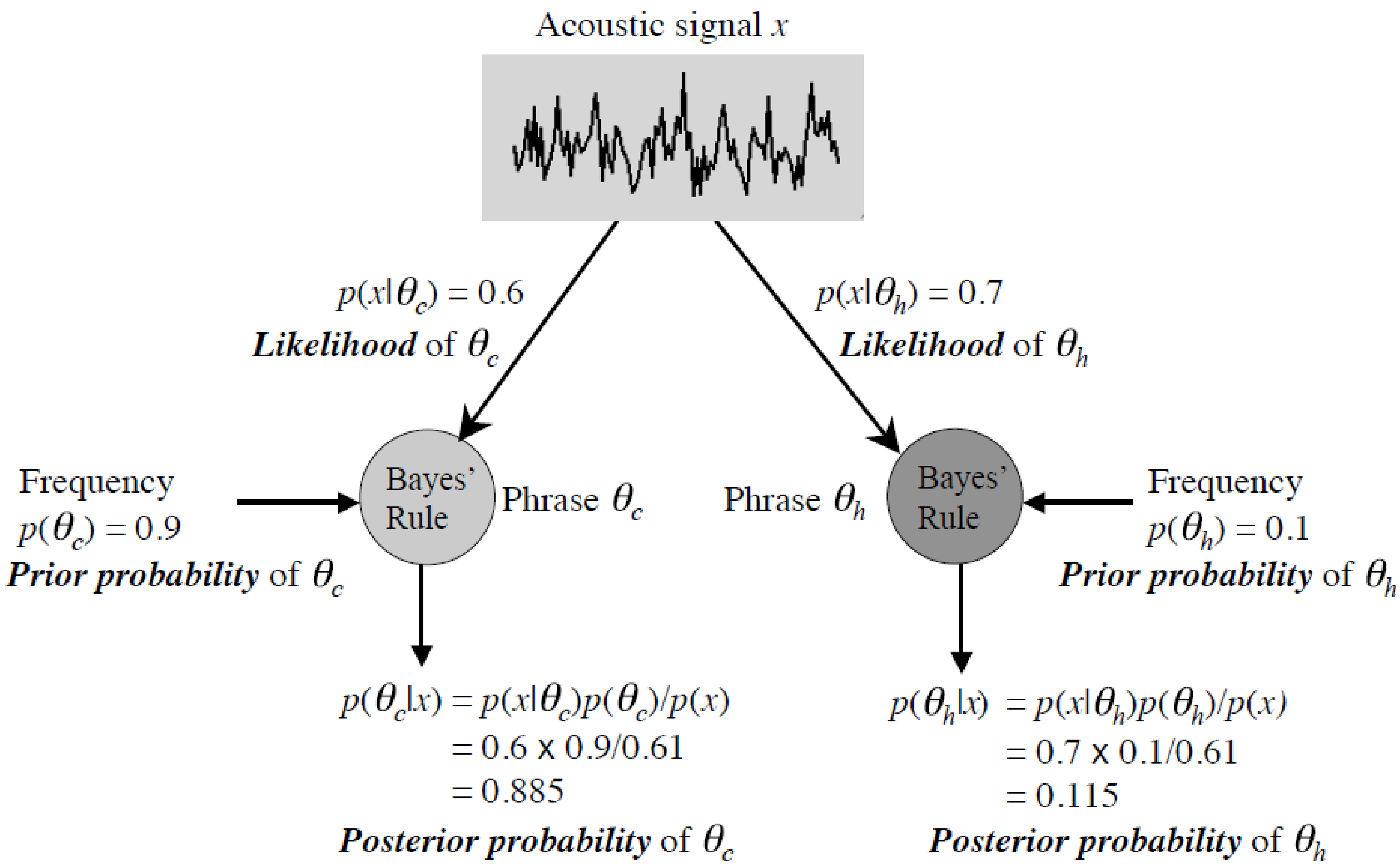
“Forkandles”



- $p(acoustic\ data|four\ candles) = p(x|\theta_c) = 0.6$
- $p(acoustic\ data|fork\ handles) = p(x|\theta_h) = 0.7$

Bayes’ Rule by J. V. Stone, 2013

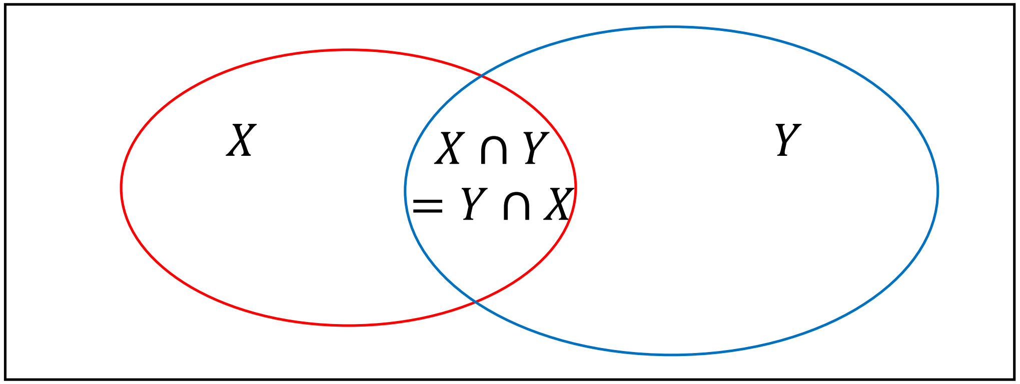
Bayesian Inference for “Forkandles”



Bayes' Rule by J. V. Stone, 2013

Bayes' Theorem

Sample Space



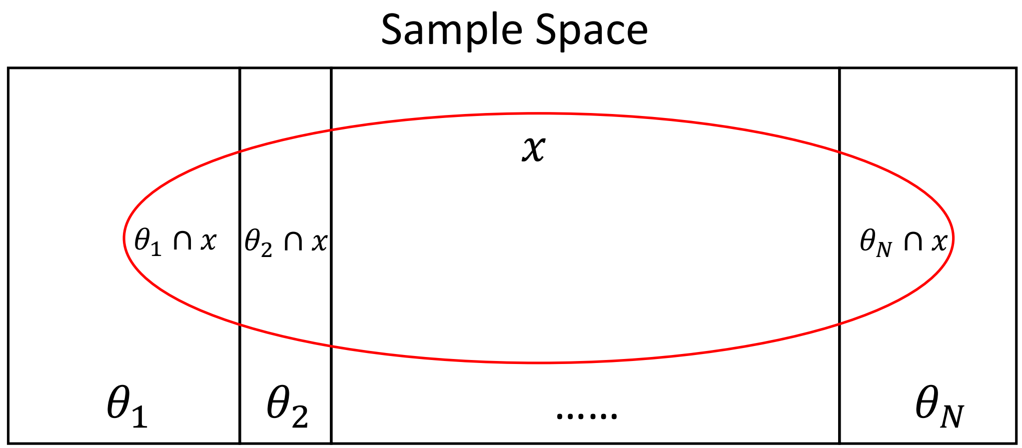
$$p(X \cap Y) = \frac{A_{X \cap Y}}{A} = \frac{A_{X \cap Y}}{A_Y}\frac{A_Y}{A} = p(X|Y)p(Y)$$

$$p(Y \cap X) = \frac{A_{Y \cap X}}{A} = \frac{A_{Y \cap X}}{A_X}\frac{A_X}{A} = p(Y|X)p(X)$$

$$p(Y|X) = \frac{p(X|Y)p(Y)}{p(X)}$$

posterior $\propto$ likelihood × prior

Bayes' Theorem



$$p(\theta_i|x) = \frac{A_{\theta_i \cap x}}{A_x} = \frac{A_{\theta_i \cap x}}{A_{\theta_i}} \frac{A_{\theta_i}}{A} \frac{A}{A_x} = p(x|\theta_i)p(\theta_i)\frac{1}{p(x)}$$

$$p(x) = p(x \cap \theta_1) + p(x \cap \theta_2) + \cdots + p(x \cap \theta_N)$$

$$p(x \cap \theta_i) = \frac{A_{x \cap \theta_i}}{A} = \frac{A_{\theta_i \cap x}}{A_{\theta_i}} \frac{A_{\theta_i}}{A} = p\left(x \middle| \theta_i\right)p(\theta_i)$$

$$p(\theta_i|x) = \frac{p(x|\theta_i)p(\theta_i)}{p(x)} = \frac{p(x|\theta_i)p(\theta_i)}{p\left(x \middle| \theta_1\right)p(\theta_1) + p\left(x \middle| \theta_2\right)p(\theta_2) + \cdots + p\left(x \middle| \theta_N\right)p(\theta_N)}$$

Bayesian Inference for Color-blind Man

$$p(color\_blind|man) = p(x|\theta) = 0.05$$

$$p(color\_blind|woman) = p(x|\theta^c) = 0.0025$$

$$p(man) = p(\theta) = 0.5$$

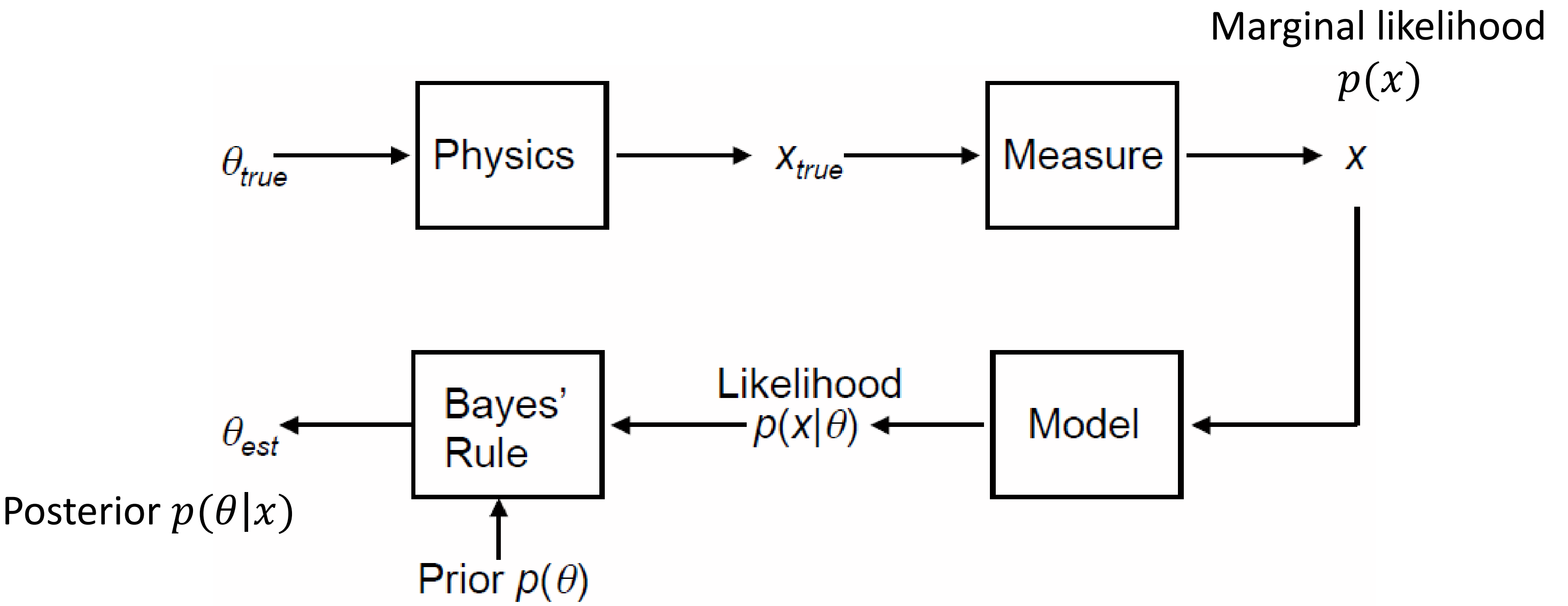
$$p(x \cap \theta) = p(x|\theta)p(\theta)$$

$$p(\theta \cap x) = p(\theta|x)p(x)$$

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)} = \frac{p(x|\theta)p(\theta)}{p(x|\theta)p(\theta) + p(x|\theta^c)p(\theta^c)} = \frac{0.05 \times 0.5}{0.05 \times 0.5 + 0.0025 \times 0.5} = 0.95238$$

Bayes' Rule by J. V. Stone, 2013

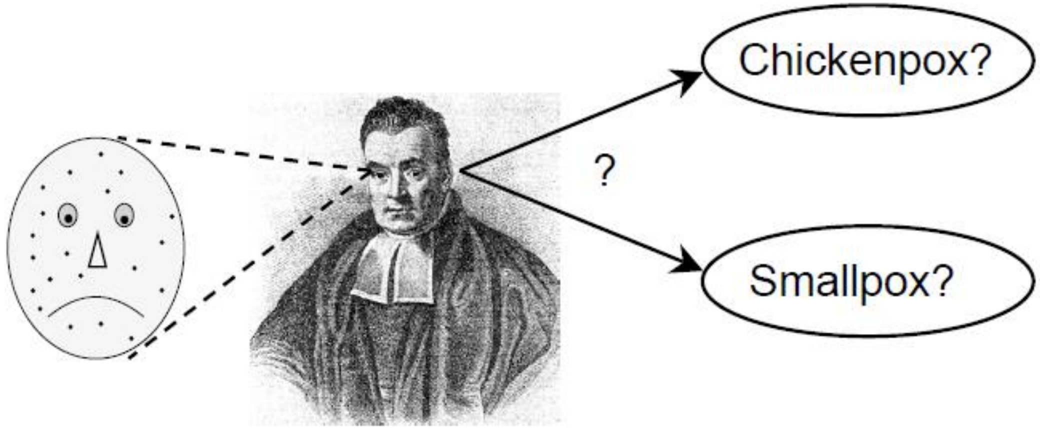
Forward and Inverse



θ	x
Hypothesis	Data (evidence)
Model	Data (evidence)
Parameter	Variable

Bayes' Rule by J. V. Stone, 2013

Diagnosing Poxy Diseases

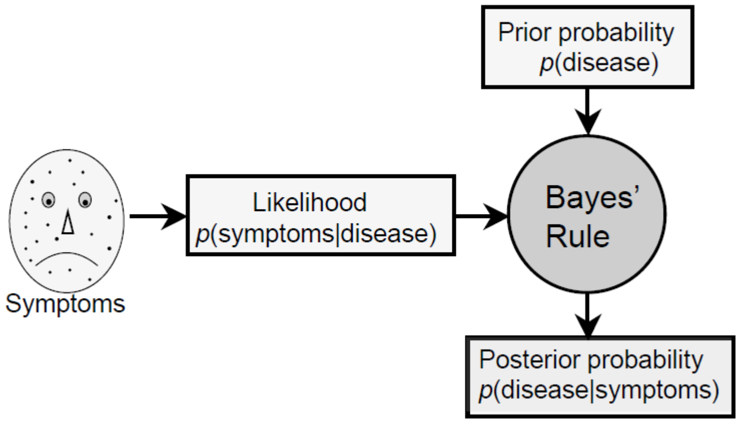


- $p(symptom|smallpox) = p(x|\theta_s) = 0.9$
- $p(symptom|chickenpox) = p(x|\theta_c) = 0.8$

- $p(smallpox|symptom) = p(\theta_s|x) = ?$
- $p(chickenpox|symptom) = p(\theta_c|x) = ?$

Bayes' Rule by J. V. Stone, 2013

Bayes' Rule for Poxy Diseases



- Likelihood : $p(symptom|disease) = p(x|\theta)$
- Prior (probability) : $p(disease) = p(\theta)$
- Posterior (probability) : $p(disease|symptom) = p(\theta|x)$

$$p(disease|symptom) = \frac{p(symptom|disease)p(disease)}{p(symptom)}$$

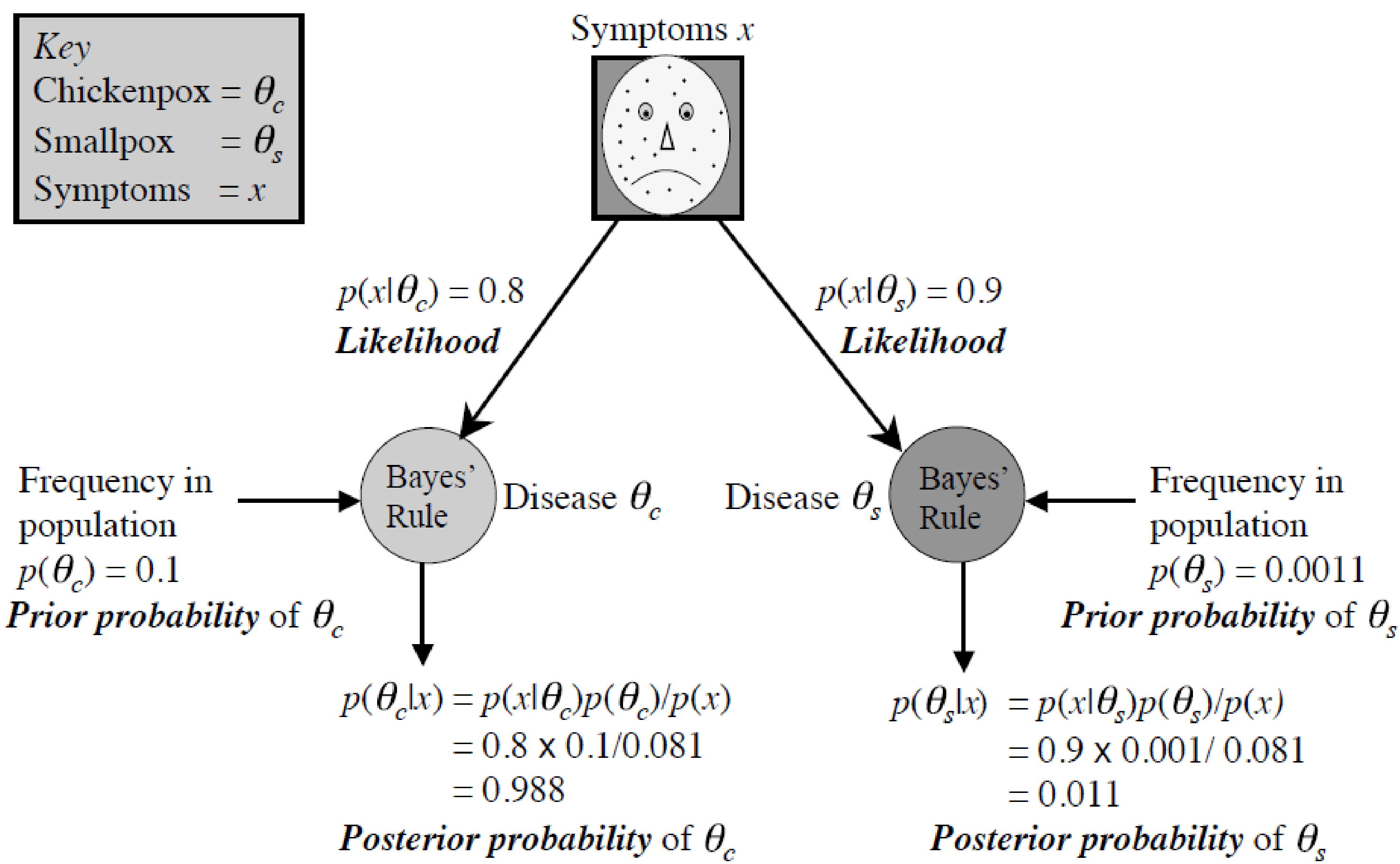
$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)}$$

$$Posterior = \frac{Likelihood \times Prior}{Marginal\ Likelihood}$$

$$p(hypothesis|data) = \frac{p(data|hypothesis)p(hypothesis)}{p(data)}$$

Bayes' Rule by J. V. Stone, 2013

Bayesian Inference for Poxy Diseases



Bayes' Rule by J. V. Stone, 2013

Bayesian Analysis of Hypothetical Sleep Apnea Diagnosis

- Probability (prevalence) of sleep apnea in a chosen population is 0.1.
- 1,000 randomly selected people were diagnosed using a medical device.

	OSA, A_+	Normal, A_-	Sum
Positive, D_+	80	90	170
Negative, D_-	20	810	830
	100	900	1,000

- A_+ : presence of sleep apnea, A_- : absence of sleep apnea, D_+ : positive diagnosis, and D_- : negative diagnosis

- Prevalence : $p(A_+) = 100/1{,}000 = 0.1$
- Sensitivity : $p(D_+|A_+) = 80/100 = 0.8$
- Specificity : $p(D_-|A_-) = 810/900 = 0.9$
- Positive predictive value (PPV) : $p(A_+|D_+) = 80/170 = 0.47$
- Negative predictive value (NPV) : $p(A_-|D_-) = 810/830 = 0.976$

$$p(A_+|D_+) = \frac{p(D_+|A_+)p(A_+)}{p(D_+|A_+)p(A_+) + p(D_+|A_-)p(A_-)} = \frac{sens \times prev}{sens \times prev + (1 - spec) \times (1 - prev)}$$

Bayesian Analysis of Hypothetical Cancer Diagnosis using MRI

- Cancer coding using MRI

1	2	3	4	5
Not cancer	Probably not cancer	Inconclusive	Probably cancer	Cancer

- Assessment data

	1	2	3	4	5	Sum
Malignant, M	7	13	22	45	91	178
Benign, B	78	56	60	5	2	201
Sum	85	69	82	50	93	379

- Sensitivity : $p(4 \cup 5|M) = (45 + 91)/178 = 0.764$
- Specificity : $p(1 \cup 2 \cup 3|B) = (78 + 56 + 60)/201 = 0.9652$

- Sensitivity and specificity of cancer diagnosis for each and higher scores

	1	2	3	4	5	Sum
Sensitivity*	1	0.96	0.89	0.76	0.51	
Specificity*	0	0.39	0.67	0.97	0.99	

EOD