Bayes Days 2000 at LANL

Three-Day Minicourse on Bayesian Analysis in Physics

Lectures presented by

Prof. Volker Dose
Max Planck Institute for Plasma Physics

Sponsored by
Enhanced Surveillance Program, Los Alamos National Laboratory

For more information, look on the web:
http://public.lanl.gov/kmh/course/BD2000.html

Organized by Ken Hanson, DX-3, 505-667-1402, kmh@lanl.gov

Bayesian Principles

V. Dose

Los Alamos Nat. Lab
April 3 – 5, 2000

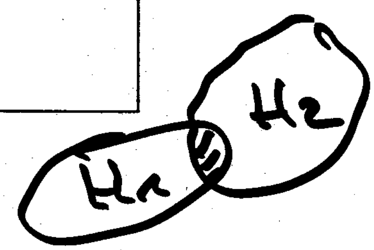
The probability axioms

The product rule (axiom I)

$$\begin{aligned} p(H, D|I) &= p(H|I)p(D|H, I) \\ &= p(D|I)p(H|D, I) \end{aligned}$$

Bayes theorem

$$\boxed{p(H|D, I) = p(H|I)\frac{p(D|H, I)}{p(D|I)}}$$



the sum rule (axiom II)

$$p(H_1 + H_2|I) = p(H_1|I) + p(H_2|I) - p(H_1, H_2|I)$$

for H_i mutually exclusive and exhaustive

$$p(\sum H_i|I) = 1$$

marginalisation

$$\sum_i p(D,H_i|I) \underset{\text{sum rule}}{=} p(D, \sum_i H_i|I) \underset{\text{product rule}}{=} p(D|I) \cdot \underbrace{p(\sum_i H_i|D, I)}_{\equiv 1}$$

$$\boxed{p(D|I) = \int dH \; p(D, H|I)}$$

Example: sum and product rule, Bayes theorem

Urn with w und b balls with masses m und M.

α, β	#	$P(\alpha, \beta)$
w, m	100	0.1
w, M	200	0.2
b, m	300	0.3
b, M	400	0.4
	—	—
	1000	1.0

$$\begin{aligned} P(w) &= P(w,m) + P(w,M) \\ &= 0.1 + 0.2 \quad = \quad 0.3 \end{aligned}$$

$$\begin{aligned} P(M|w) &= \frac{P(M,w)}{P(w)} \\ &= \frac{P(M,w)}{P(m,w) + P(M,w)} \\ &= \frac{0.2}{0.1 + 0.2} \quad = \quad \frac{2}{3} \end{aligned}$$

Assigning probabilities

(A) The principle of maximum entropy

$$S = -\int p(x)\ \ln p(x)\ dx$$

testable information

$$M^g(p) = \int g(x) p(x) dx$$

$$\partial\Phi = \partial \left\{ -\int p(x) \ln p(x) + \lambda \left(M^g - \int g(x) p(x) dx \right) \right\}$$
$$= \int \partial p \left\{ -\ln p(x) - 1 - g(x)\lambda \right\} \overset{!}{=} 0$$

solution: $\boxed{p(x) = p_0 \exp\{-\lambda g(x)\}}$

special cases: $(g(x) = 1, M^g = 1)$, $(g(x) = x, M^g = \mu)$

$$p(x) = \exp\{-x/\mu\}/\mu\ , \qquad 0 \le x \le \infty$$

$(g(x){=}1, M^g{=}1)$, $(g(x){=}x, M^g{=}\mu)$, $(g(x){=}x^2, M^g{=}\sigma^2{+}\mu^2)$

$$p(x) = \exp\left\{ -\frac{1}{2\sigma^2}(x-\mu)^2 \right\} / \sigma\sqrt{2\pi}, \qquad -\infty \le x \le \infty$$

Assigning probabilities

(B) Transformation invariance

$$p(x|I)dx = p(y|I)dy$$

1.) let x be a location variable

transform $y = x + b$

$$p(x|I)\ dx = p(x+b|I)\ dx \qquad \forall\ b$$

$$\boxed{p(x|I) = const \qquad -\infty < x < \infty}$$

$$p(x|I) = 1/2B, \qquad B \to \infty$$

2.) let x be a scale variable

transform $y = \alpha x,\ dy = \alpha dx$

$$p(x|I)\ dx = \alpha\ p(\alpha x|I)\ dx$$ Jeffreys prior

$$\boxed{p(x|I) = 1/x \qquad 0 < x < \infty}$$

$\frac{1}{x+\varepsilon} e^{-\mu x}$ | $\frac{1}{B} < x < B \qquad B \to \infty$

$d_i - a x_i = \varepsilon_i, \quad \langle \varepsilon_i \rangle = 0, \quad \langle \varepsilon_i^2 \rangle = \sigma^2$

Marginalization: An example

$$p(\vec{d}|a, \vec{x}, \sigma, I) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N \exp\left\{-\frac{1}{2\sigma^2}\sum_i (d_i - a x_i)^2\right\}$$

$$\begin{aligned} p(\vec{d}|\vec{x}, \sigma, I) &= \int p(\vec{d}, a|\vec{x}, \sigma, I)\, da \\ &= \int p(a|\not{\vec{x}}, \not{\sigma}, I)\, p(\vec{d}|a, \vec{x}, \sigma, I)\, da \end{aligned}$$

choose flat prior for a

$$p(a|I) = \frac{1}{2A}, \qquad -A \leq a \leq A$$

$$p(\vec{d}|\vec{x}, \sigma, I) = \frac{1}{2A}\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N \int_{-A}^{A} \exp\left\{-\frac{1}{2\sigma^2}\sum_i (d_i - a x_i)^2\right\} da$$

define

$$\sum d_i = N\overline{d}, \ \sum x_i d_i = N\overline{xd}, \ \sum d_i^2 = N\overline{d^2}, \ \sum x_i^2 = N\overline{x^2}$$

$$\begin{aligned} p(\vec{d}|\vec{x}, \sigma, I) &= \frac{1}{2A}\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N \left\{\frac{2\pi\sigma^2}{N\overline{x^2}}\right\}^{\frac{1}{2}} \\ &\quad \cdot \exp\left\{-\frac{N\overline{d^2}}{2\sigma^2}\left(1 - \frac{(\overline{xd})^2}{\overline{x^2}\ \overline{d^2}}\right)\right\} \end{aligned}$$

↑ a2

Product rule: An illustration

$$p(x,y) = \frac{1}{Z}\text{exp}\left\{-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2} - \frac{2axy}{2\sigma_x\sigma_y}\right\}, \quad \frac{1}{Z} = \frac{\sqrt{1-a^2}}{2\pi\sigma_x\sigma_y}$$

$$0 \le a < 1$$

product rule

$$p(x,y) = p(x)p(y|x)$$

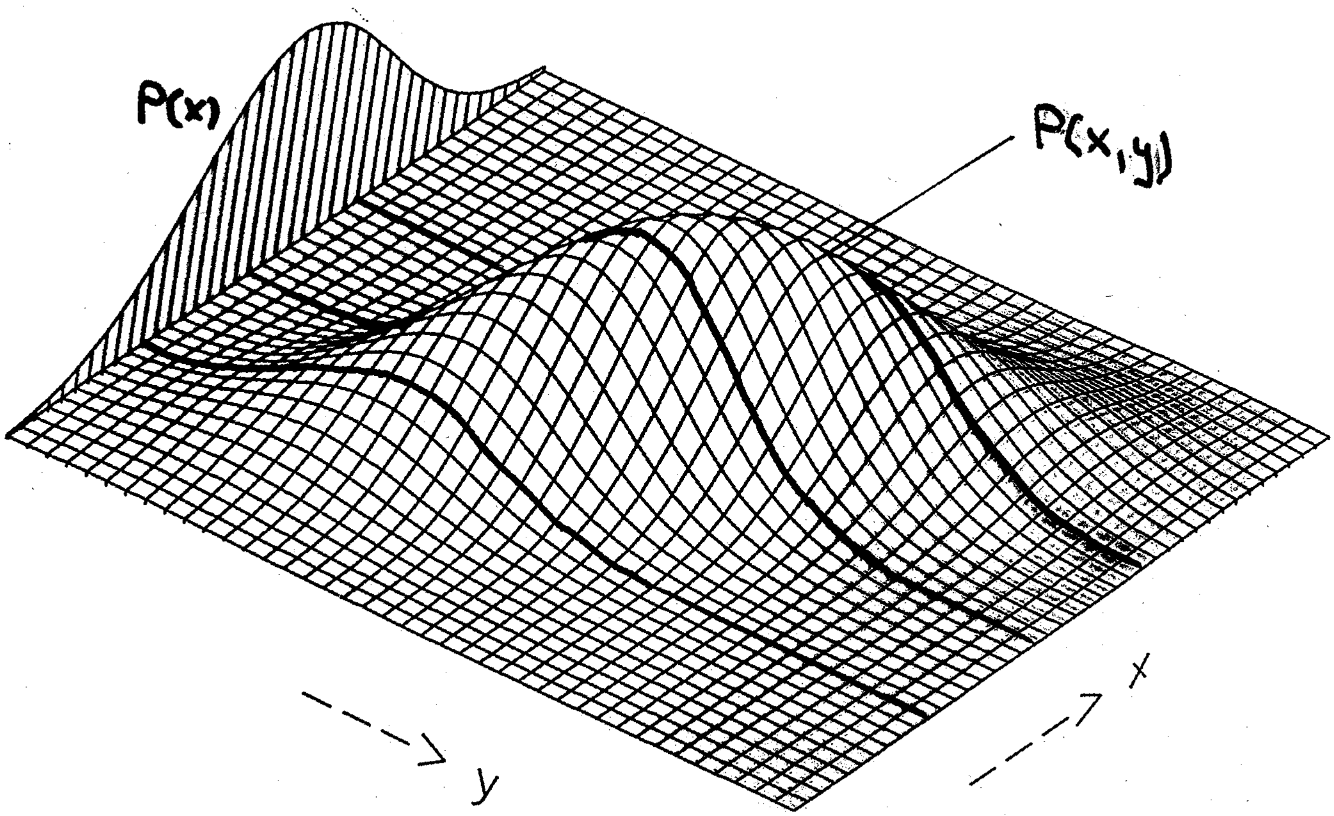
sum rule

$$p(x) = \int p(x,y)dy = \frac{1}{Z}\sigma_y\sqrt{2\pi}\text{exp}\left\{-\frac{x^2}{2\sigma_x^2}(1-a^2)\right\}$$

$$p(y|x) = \frac{p(x,y)}{p(x)}$$

$$p(y|x) = \frac{1}{\sigma_y\sqrt{2\pi}}\text{exp}\left\{-\frac{1}{2\sigma_y^2}\left(y + ax\frac{\sigma_y}{\sigma_x}\right)^2\right\}$$

$p(x,y)$ = series of shifted Gaussians $p(y|x)$
with amplitude given by $p(x)$


P(x)
P(x,y)
y
x

Bayesian model comparison

$$p(\vec{d}|a,\vec{x},\sigma,M_1) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(d_i - a x_i)^2\right\}$$

$$p(\vec{d}|a,\vec{x},\sigma,M_2) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(d_i - a \underbrace{(e^{x_i} - 1)}_{x_i^*})^2\right\}$$

Sought

$$p(M_i|\vec{d},\vec{x},\sigma)$$

Bayes Theorem

$$p(M_i|\vec{d},\vec{x},\sigma) = \frac{p(M_i|\vec{x},\sigma)\cdot p(\vec{d}|\vec{x},\sigma,M_i)}{p(\vec{d}|\vec{x},\sigma)}$$

$$\frac{p(M_i|\vec{d},\vec{x},\sigma)}{p(M_k|\vec{d},\vec{x},\sigma)} = \underbrace{\frac{p(M_i|\vec{x},\sigma)}{p(M_k|\vec{x},\sigma)}}_{\text{prior odds}} \cdot \underbrace{\frac{p(\vec{d}|\vec{x},\sigma,M_i)}{p(\vec{d}|\vec{x},\sigma,M_k)}}_{\text{Bayes factor } B_{ik}}$$

Calibration of Bayes factor B_{10}*)

B_{10}	$\ln B_{10}$	evidence for M_1
< 1	< 0	negative (supports M_0)
$1 \cdot\cdot 3$	$0 \cdots 1$	barely worth mentioning
$3 \cdot\cdot 20$	$1 \cdots 3$	positive
$20 \cdot\cdot 150$	$3 \cdots 5$	strong
> 150	> 5	very strong

*) After A. E. Raftery in "Markov Chain Monte Carlo in Practice", edited by Gilks, Richardson and Spiegelhalter, Chapman and Hall 1996

Approximate consideration

$$p(\vec{d}|\vec{x},\sigma,M_1) = \int p(\vec{\theta}|M_1)\cdot p(\vec{d}|\vec{x},\sigma,\vec{\theta},M_1)d\vec{\theta}$$

$$\approx p(\widehat{\vec{\theta}}|M_1)\cdot\int d\vec{\theta}\ \mathcal{L}(\vec{\theta})$$

$$\approx p(\widehat{\vec{\theta}}|M_1)\cdot\mathcal{L}(\widehat{\theta})(\partial\theta)^{E_1}$$

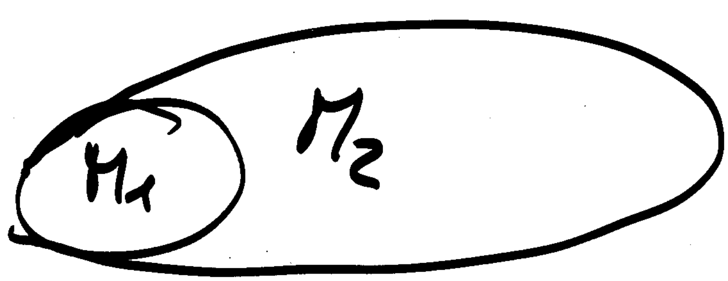
$$\int p(\vec{\theta}|M_1)d\vec{\theta} = 1 = p(\widehat{\vec{\theta}}|M_1)\cdot(\Delta\theta)^{E_1}$$

$$p(\vec{d}|\vec{x},\sigma,M_1)\approx\mathcal{L}(\widehat{\theta})\left(\frac{\partial\theta}{\Delta\theta}\right)^{E_1}$$

$$B_{12} = \frac{\mathcal{L}(\widehat{\theta},M_2)}{\mathcal{L}(\widehat{\theta},M_1)}\cdot\left(\frac{\partial\theta}{\Delta\theta}\right)^{E_2-E_1}$$

$\gtrsim 1$ $\quad < 1$

Occam's factor



Plasma Physics and Controlled Fusion

Reference number: **102991** Surname: **Fischer**

This service enables authors to track the current status and progress of their article. If you have any questions, please contact the Editorial Office at the usual address, ppcf@ioppublishing.co.uk

However please note that the Office is not able to give details of referees' recommendations until a final decision has been taken.

Electron energy distribution reconstruction in low-pressure helium plasmas from optical measurements

Date	Status
12/07/99	Referee 8's report received
25/06/99	Article sent to Board Member (referee 8)
24/06/99	Referee 7 could not report
02/06/99	Article sent to referee 7
02/06/99	Referee 5 did not reply to repeated reminders
01/06/99	Referee 6's report received
25/05/99	Referee 5 reminded again
18/05/99	Referee 5 reminded
17/05/99	Referee 5 reminded
06/05/99	Article sent to referee 6
05/05/99	Referee 4 [redacted]
29/04/99	Article sent to referee 5
29/04/99	Referee 3 [redacted] but recommended alternative
28/04/99	Article sent to referee 4
26/04/99	Referee 2 [redacted]
16/04/99	Article sent to referee 3
13/04/99	Referee 1 [redacted]
30/03/99	Article sent to referee 2
30/03/99	Article sent to referee 1
25/03/99	Article received by IOPP Office
25/03/99	Electronic file received by IOPP Office

1st Referee's Report

The paper deals with an elegant technique based on Bayesan data analysis for the deconvolution of eedf from optical measurements.

The paper can be published after small revision taking into account these minor comments:

1) the most extensive calculations of eedf in Helium are reported By Capriati et al. Plasma Chemistry and Plasma Processing 12,237(1992)
2) It should be interesting to compare for non local effects with the measurements of Dilecce et al?J. Appl. Phys., 69 (1991) 121-128.for RF plasmas
3) The authors should better report about their Langmuir probe measurements of eedf

Board Member's Report

This is an interesting paper which merits publication with some very minor revisions. The authors have utilized Bayesian statistics in an ingenious way to determine the most probable electron energy distribution function without constraining the analysis with assumptions about the functional form. This is a very nice piece of work.

I have a few suggestions for improvement in the paper. At some point, it would be helpful if the authors listed the wavelengths of the spectral lines that they used. This could, for example, be included in Table 1 with little increase in the length of the paper.

There is a point of terminology which needs to be improved to reduce confusion. At several points in the paper, the authors say that a certain feature in the distribution function is "hardly significant". See, for example, the discussion on page 14. To me, this phrase means "not significant at all". However, the context suggests that the authors are using it to mean "marginally significant" or "barely significant". The authors should improve the phrasing so that there is no possibility of confusion.