The Geometry of Imprecise Inference

Miķelis Bickis

University of Saskatchewan

ISIPTA'15
20 July 2015
Pescara, Italy

- I am Professor Emeritus in the Department of Mathematics and Statistics at the

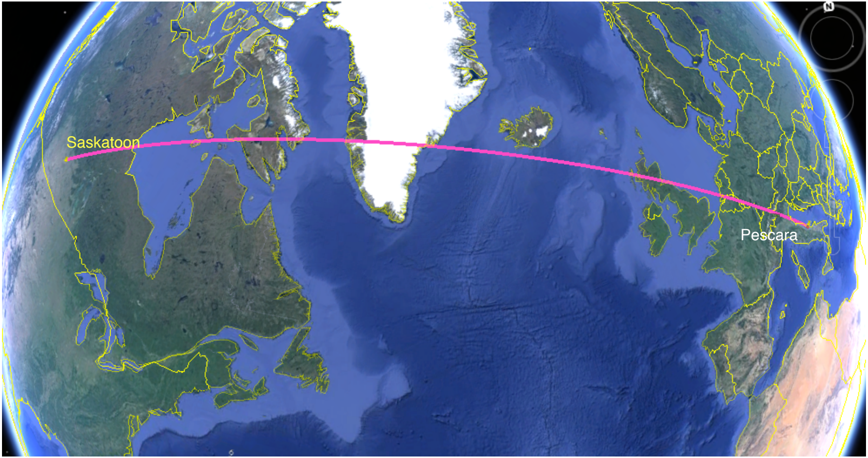
University of Saskatchewan

in Saskatoon

- My recent graduates:
  - Osama Bataineh, PhD 2012
  - Chel Hee Lee, PhD 2014
- Two current students
- Research supported by the Natural Sciences and Engineering Research Council of Canada

NSERC CRSNG

Where is Saskatoon?



Bayes' theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Bayes' theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$$\log \frac{P(B|A)}{P(A)} = \log P(B|A) - \log P(B)$$

Bayes' theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$$\log \frac{P(B|A)}{P(A)} = \log P(B|A) - \log P(B)$$

$$\log \frac{P(A|B)}{P(A^c|B)} = \log \frac{P(B|A)}{P(B|A^c)} + \log \frac{P(A)}{P(A^c)}$$

Bayes' theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$$\log \frac{P(B|A)}{P(A)} = \log P(B|A) - \log P(B)$$

$$\log \frac{P(A|B)}{P(A^c|B)} = \log \frac{P(B|A)}{P(B|A^c)} + \log \frac{P(A)}{P(A^c)}$$

$$\log \frac{P(A|B)}{P(A)} = \log \frac{P(B|A)}{P(B|A^c)} + \log \frac{P(A^c|B)}{P(A^c)}$$

Bayes' theorem:

$$\log \frac{P(A|B)}{P(A)} = \log \frac{P(B|A)}{P(B|A^c)} + \log \frac{P(A^c|B)}{P(A^c)}$$

Bayes' theorem:

$$\log \frac{P(A|B)}{P(A)} = \log \frac{P(B|A)}{P(B|A^c)} + \log \frac{P(A^c|B)}{P(A^c)}$$

$$\log \frac{d\Pi_y}{d\Pi_0}(\theta) = \log \frac{dP_\theta}{dP_0}(y) + \text{something else}$$

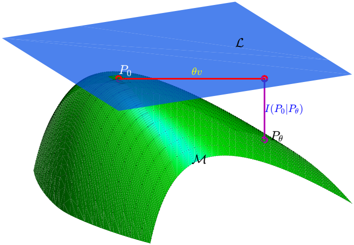
Bayes' theorem:

$$\log \frac{P(A|B)}{P(A)} = \log \frac{P(B|A)}{P(B|A^c)} + \log \frac{P(A^c|B)}{P(A^c)}$$

$$\log \frac{d\Pi_y}{d\Pi_0}(\theta) = \log \frac{dP_\theta}{dP_0}(y) + \text{something else}$$

$$\log \frac{d\Pi_y}{d\Pi_0}(\boldsymbol{\theta}) = \boldsymbol{\theta}^\top \mathbf{v}(y) - I(P_0|P_{\boldsymbol{\theta}}) - \psi(y)$$

An **exponential family of distributions** can be represented in terms of the vector space of minimal sufficient statistics (i.e. functions on the observation space.)
The manifold of distributions $\mathcal{M}$ maps uniquely unto a tangent space $\mathcal{L}$. spanned by the minimal sufficient statistics.

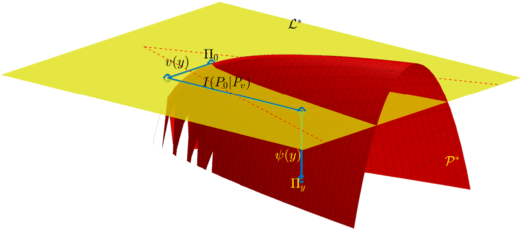


$$\log \frac{dP_\theta}{dP_0} = \theta v - I(P_0|P_\theta)$$

A prior distribution is a distribution on $\mathcal{L}$. An exponential family of priors can be expressed in terms of the space of linear functions on $\mathcal{L}$.

This dual space will include the evaluation functional $v \mapsto v(y)$, and thus will include all possible posteriors from any prior in the family.

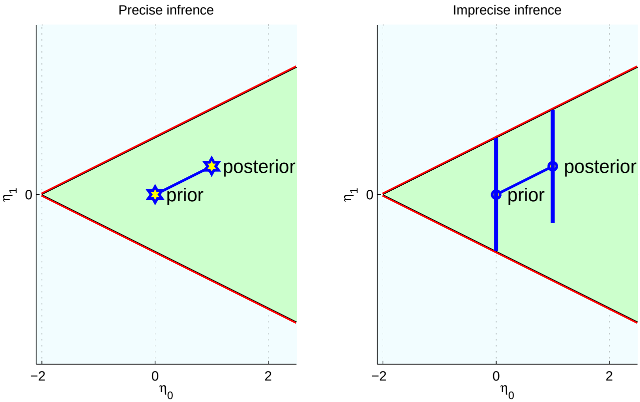
$$\log \frac{d\Pi_y}{d\Pi_0}(v) = v(y) - I(P_0|P_v) - \psi(y)$$



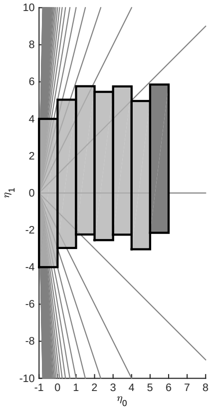
The lower envelope theorem says that a lower prevision can be expressed as the infimum over a set of linear previsions.
This means that an imprecise prior can represented by a set of precise priors.
In the expression

$$\log \frac{d\Pi_y}{d\Pi_0}(\boldsymbol{\theta}) = \boldsymbol{\theta}^\intercal \mathbf{v}(y) - I(P_0|P_{\boldsymbol{\theta}}) - \psi(y)$$

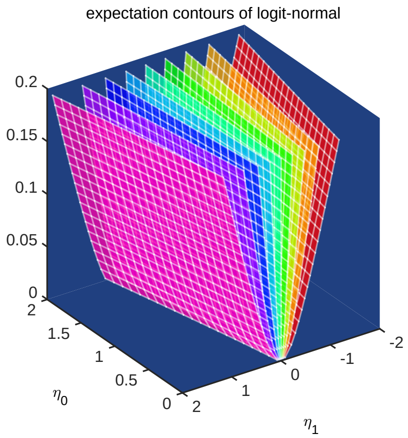
the first two terms on the right do not depend on the prior, and thus would translate all priors in the same way. The third term does not depend on the model parameter $\boldsymbol{\theta}$.


Precise infrence
posterior
prior
η_1
0
−2
0
2
η_0
Imprecise infrence
posterior
prior
η_1
0
−2
0
2
η_0

Normal family with known variance.



Binomial model with normal prior of logit



Censored exponential model with gamma prior