A Bayesian approach to informing decision makers: Comparisons to minimizing relative entropy

The Ninth Workshop on Information Theoretic Methods in Science and Engineering

19 September 2016

David Bickel
University of Ottawa

Bayes's theorem

$$\frac{P(\theta \in \mathcal{H} | X = x)}{P(\theta \notin \mathcal{H} | X = x)} = \frac{f(x | \theta \in \mathcal{H})}{f(x | \theta \notin \mathcal{H})} \frac{P(\theta \in \mathcal{H})}{P(\theta \notin \mathcal{H})}$$

- posterior odds = Bayes factor × prior odds
- Bayes factor = strength of evidence

Decision-maker's loss function

$$\ell_c\left(h, \widehat{h}\right) = \begin{cases} 0 & \text{if } \widehat{h} = h \\ c & \text{if } \widehat{h} = 1, h = 0 \\ 1 & \text{if } \widehat{h} = 0, h = 1 \end{cases}$$

Bayes action

$$\widehat{h}_c\left(\frac{\pi_1}{\pi_0}\right) = \arg\min_{\widehat{h}=0,1}\left(\pi_0 \ell_c\left(0,\widehat{h}\right) + \pi_1 \ell_c\left(1,\widehat{h}\right)\right)$$

$$= \begin{cases} 0 & \text{if } \pi_1/\pi_0 < c \\ 1 & \text{if } \pi_1/\pi_0 > c \end{cases}$$

Scientist's regret function

$$\rho_{c,\kappa}(\omega,\widehat{\omega}) = \begin{cases} 0 & \text{if } \widehat{h}_c(\omega) = \widehat{h}_c(\widehat{\omega}) \\ 1 & \text{if } \widehat{h}_c(\omega) = 1, \widehat{h}_c(\widehat{\omega}) = 0 \\ \kappa & \text{if } \widehat{h}_c(\omega) = 0, \widehat{h}_c(\widehat{\omega}) = 1 \end{cases}$$

Scientist's odds to report

$$\rho_{(\underline{c},\overline{c}),\kappa}\left(\omega,\widehat{\omega}\right) \;=\; \int_{\underline{c}}^{\overline{c}} \rho_{c,\kappa}\left(\omega,\widehat{\omega}\right) q\left(c\right) dc$$

$$\rho_{\kappa}\left(\omega,\widehat{\omega}\right) = \lim_{\underline{c}\to 0} \rho_{(\underline{c},1/\underline{c}),\kappa}\left(\omega,\widehat{\omega}\right)$$

$$\omega_{\Omega_1,\kappa} = \arg \inf_{\widehat{\omega}\in[0,\infty]} \sup_{\omega\in\Omega_1} \rho_{\kappa}\left(\omega,\widehat{\omega}\right)$$

Known cost ratio

$$\begin{aligned}\rho_\kappa(\omega,\widehat{\omega}) &= \lim_{\underline{c}\to 0}\int_{\underline{c}}^{1/\underline{c}} \rho_{c,\kappa}(\omega,\widehat{\omega})\,\delta(c-c')\,dc \\ &= \rho_{c',\kappa}(\omega,\widehat{\omega})\end{aligned}$$

$$\widehat{\Omega}_{1,c'} = \begin{cases} [0,c'[& \text{if } \kappa > 1 \\ ]c',\infty] & \text{if } \kappa < 1 \end{cases}$$

Unknown cost ratio

$$q\left(bc\right) = k\left(b\right) q\left(c\right)$$

$$\Rightarrow \quad q\left(c\right) \propto 1/c$$

$$C \sim q \quad \left(1/C\right) \sim q$$

- Scale & reciprocal invariance ⇒ Benford's law

Why log odds?

$$\rho_\kappa(\omega, \widehat{\omega}) \propto \begin{cases} \log \omega - \log \widehat{\omega} & \text{if } \widehat{\omega} < \omega \\ 0 & \text{if } \widehat{\omega} = \omega \\ (\log \widehat{\omega} - \log \omega)\,\kappa & \text{if } \widehat{\omega} > \omega \end{cases}$$

Geometric mean odds

$$\omega_{\Omega_1,\kappa} = \arg \inf_{\widehat{\omega}\in[0,\infty]} \sup_{\omega\in\Omega_1} \rho_\kappa\left(\omega, \widehat{\omega}\right)$$

$$\omega_{\Omega_1,\kappa} = \left(\underline{\omega}^\kappa \overline{\omega}\right)^{\frac{1}{\kappa+1}}$$

Geometric mean Bayes factor

$$B_{\mathcal{B}}\left(\kappa\right) = \arg \inf_{\widehat{B} \in [0,\infty]} \sup_{B \in \mathcal{B}} \rho_{\kappa}\left(\frac{\pi_1}{1-\pi_1}B, \frac{\pi_1}{1-\pi_1}\widehat{B}\right)$$

$$B_{\mathcal{B}}\left(\kappa\right) = \left(\underline{B}^{\kappa}\overline{B}\right)^{\frac{1}{\kappa+1}}$$

$$B_{\mathcal{B}}\left(0\right) = \overline{B} \qquad B_{\mathcal{B}}\left(1\right) = \sqrt{\underline{B}\overline{B}} \qquad B_{\mathcal{B}}\left(\infty\right) = \underline{B}$$

Evidential distance

$$\rho_\kappa\left(P,\widehat{P}\right)=\sup_{\mathcal{H}\in\mathfrak{H}}\rho_\kappa\left(\frac{P\left(\mathcal{H}\right)}{1-P\left(\mathcal{H}\right)},\frac{\widehat{P}\left(\mathcal{H}\right)}{1-\widehat{P}\left(\mathcal{H}\right)}\right)$$

$$\rho\left(P,\widehat{P}\right)=\sup_{\mathcal{H}\in\mathfrak{H}}\left|\log\left(\frac{P\left(\mathcal{H}\right)}{1-P\left(\mathcal{H}\right)}\Big/\frac{\widehat{P}\left(\mathcal{H}\right)}{1-\widehat{P}\left(\mathcal{H}\right)}\right)\right|$$

Minimax evidential distance or minimax redundancy?

$$P_{\Gamma,\kappa} = \arg\inf_{\widehat{P}\in\mathcal{P}} \sup_{P\in\Gamma} \rho_\kappa\left(P,\widehat{P}\right)$$

$$P_{\Gamma,\kappa} = P_\Gamma \qquad P_\Gamma = \arg\inf_{\widehat{P}\in\mathcal{P}} \sup_{P\in\Gamma} \rho\left(P,\widehat{P}\right)$$

$$\omega_{\Omega_1(\Gamma),1} = P_\Gamma\left(\{1\}\right)/P_\Gamma\left(\{0\}\right) \qquad \left(P_\Gamma\right)^x = P_{\Gamma(x)}$$

Minimum evidential distance or maximum entropy?

$$P_{\Gamma,\kappa,P_0} = \arg\inf{}_{\widehat{P}\in\Gamma}\,\rho_\kappa\left(P_0,\widehat{P}\right)$$

$$P_{\Gamma,P_0} = \arg\inf_{\widehat{P}\in\Gamma}\rho\left(P_0,\widehat{P}\right)$$

$$P_{\Gamma,\kappa,P_0} = P_{\Gamma,P_0}$$

$$\left(P_{\Gamma,P_0}\right)^x = P_{\Gamma(x),P_0^x}$$

Finally

Preprint coming soon … davidbickel.com

"Reporting Bayes factors or probabilities to decision makers of unknown loss functions"

NSERC CRSNG

Funded by the Natural Sciences and Engineering Research Council of Canada

$$P_\Gamma = \arg \inf_{\widehat{P} \in \mathcal{P}} \sup_{P \in \Gamma} \rho\left(P, \widehat{P}\right) \qquad P_{\Gamma, P_0} = \arg \inf_{\widehat{P} \in \Gamma} \rho\left(P_0, \widehat{P}\right)$$