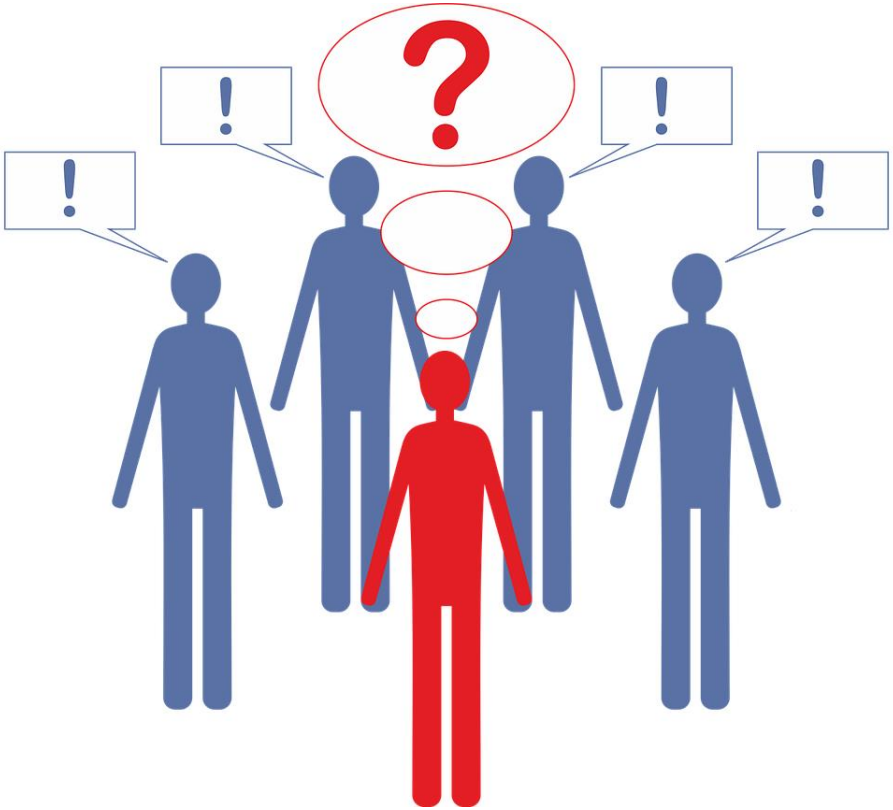
Using the Data Agreement Criterion to Rank Experts' Beliefs



Universiteit Utrecht

Slides available at: DucoVeen.com

profitwise

Data & Expert Knowledge

- Who best predicts new data?
- Do my data and experts agree?
- Are my data reliable?

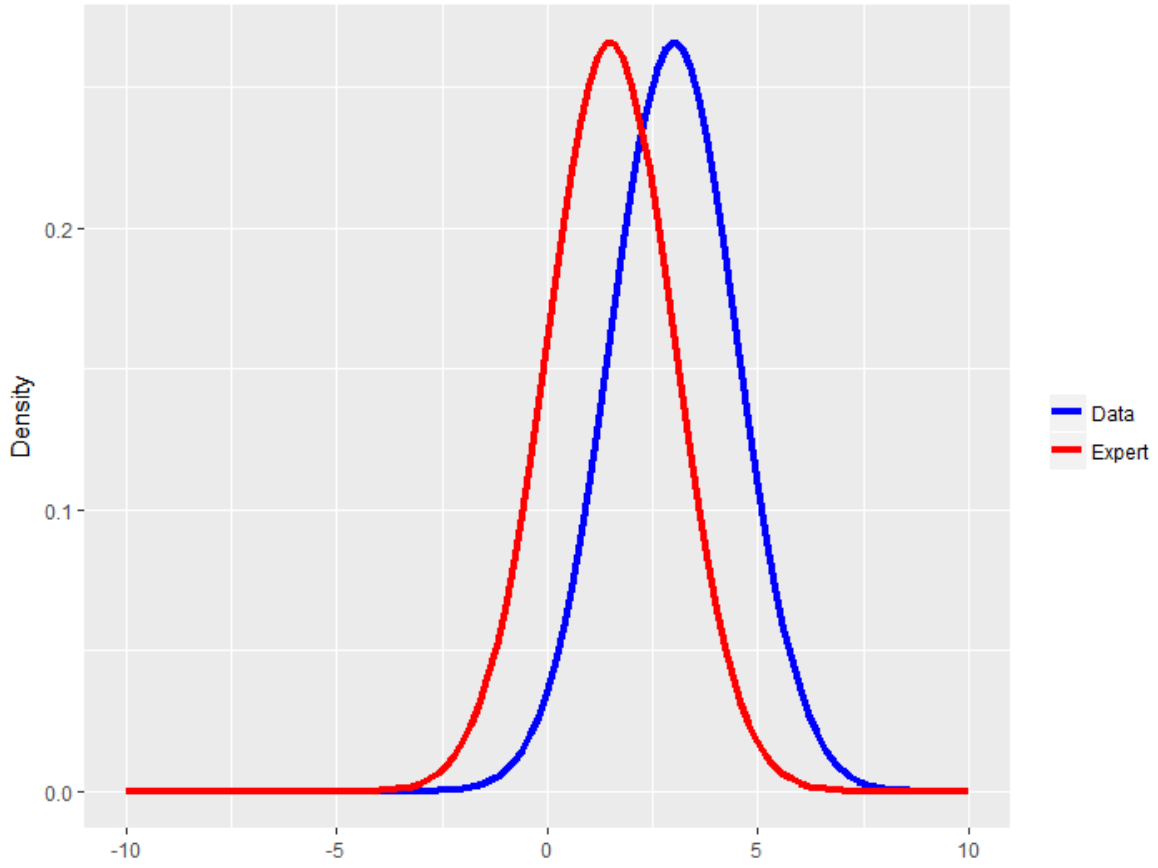
Selecting a Measure

- Evidence based evaluation of experts' beliefs
- Including knowledge and uncertainty
- Allow ranking to asses relative merit

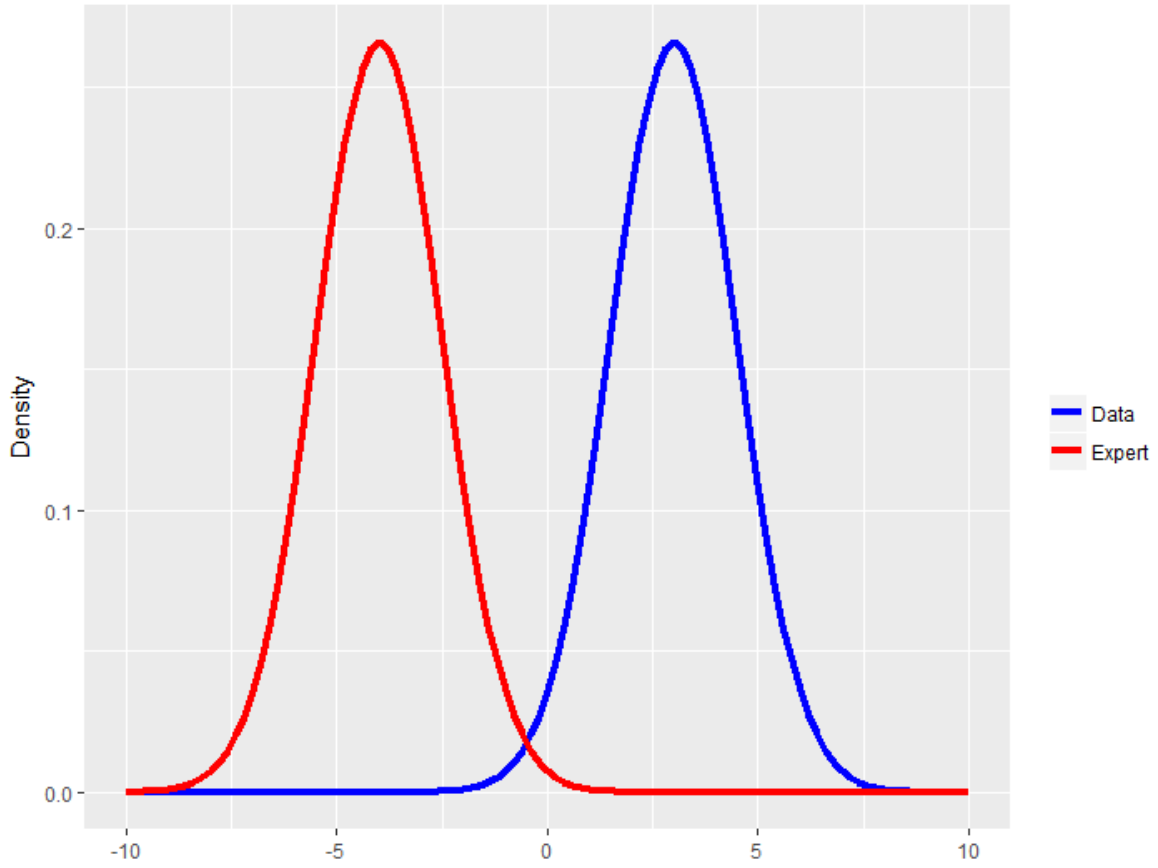
Selecting a Measure

- Represent beliefs in probability distributions
  - Prior belief
- Collect data
- Measure discrepancy
  - Prior-data Conflict

Prior-data Agreement



Prior-data Conflict



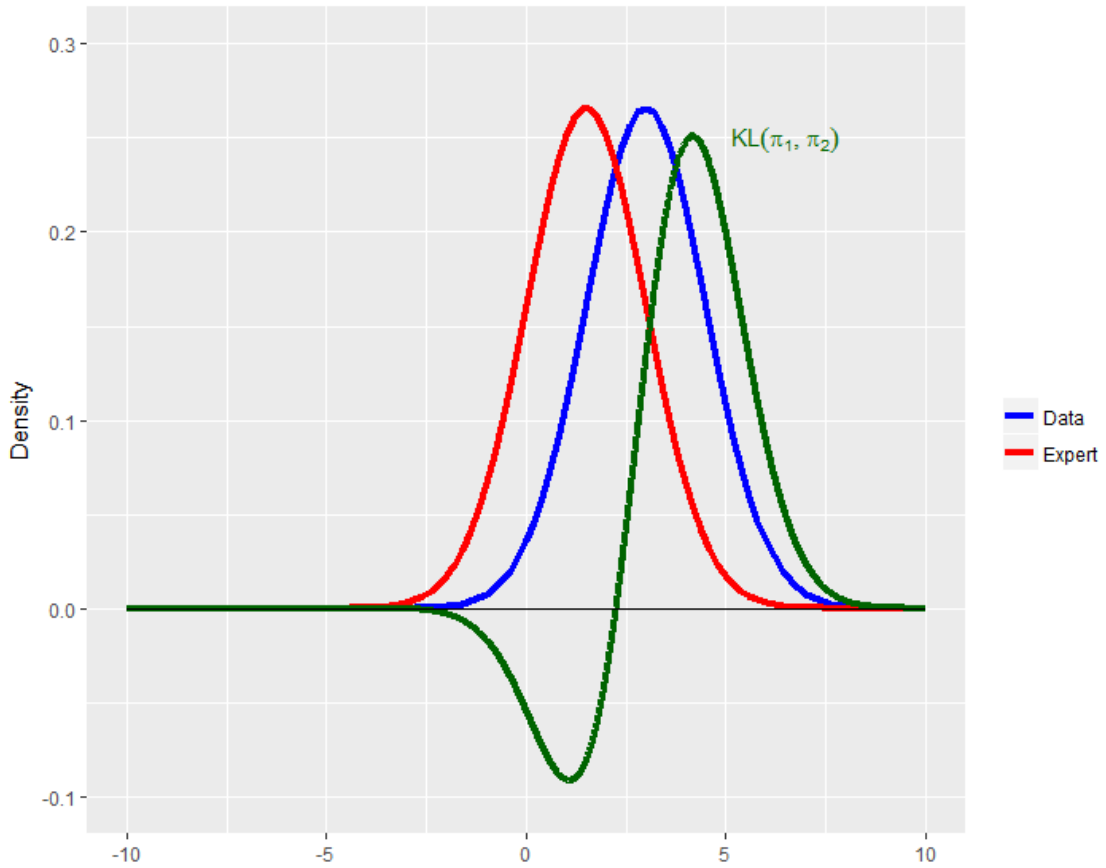
Data Agreement Criterion[1]

- Ratio of two Kullback-Leibler divergences[2]

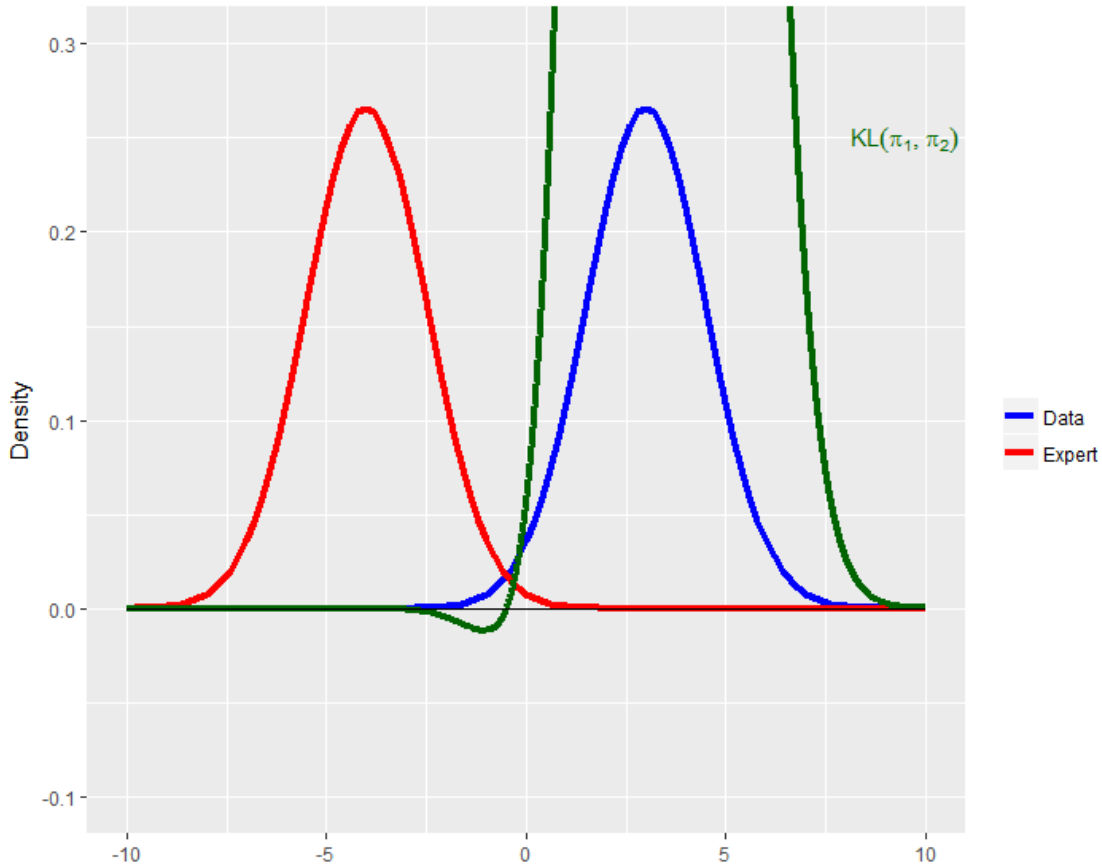
$$KL\left(\pi_1 || \pi_2\right) = \int_{\Theta} \pi_1\left(\theta\right) log \frac{\pi_1\left(\theta\right)}{\pi_2\left(\theta\right)} d\theta$$

- Measures loss of information

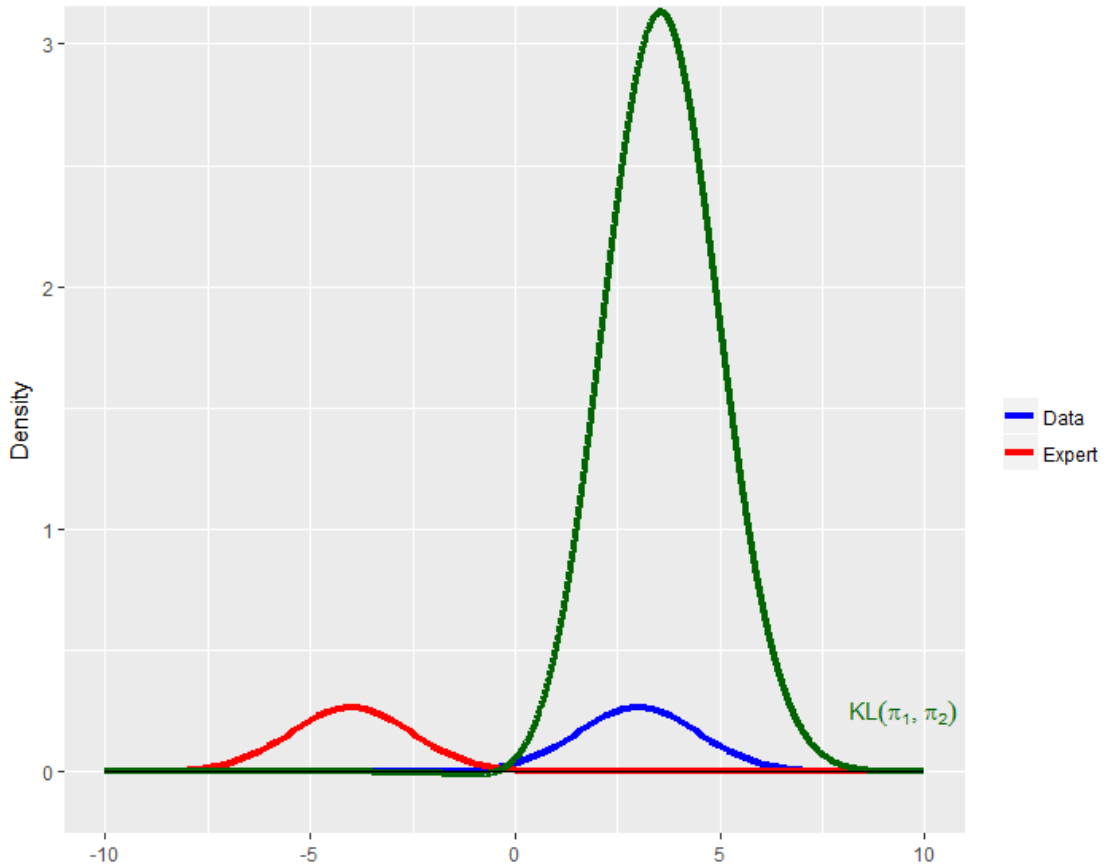
Kullback-Leibler Divergence



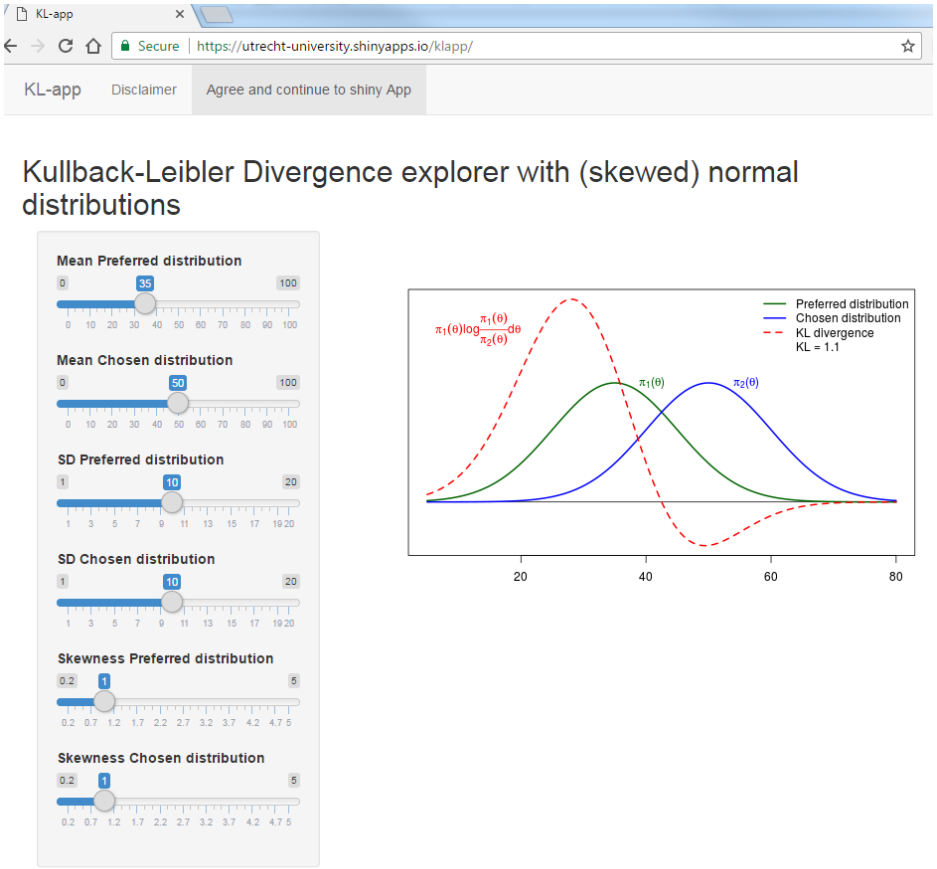
Kullback-Leibler Divergence



Kullback-Leibler Divergence



Kullback-Leibler App[2]



Data Agreement Criterion

- Preferred distribution $\pi^J(\theta|\mathbf{y})$
  - Posterior Data & Benchmark prior $\pi^J(\theta)$
- Loss of information for choosing expert's beliefs $\pi(\theta)$
- Loss of information for choosing benchmark prior $\pi^J(\theta)$

$$\mathrm{DAC} = \frac{KL\left[\pi^{\mathrm{J}}\left(\theta|\mathbf{y}\right)||\pi\left(\theta\right)\right]}{KL\left[\pi^{\mathrm{J}}\left(\theta|\mathbf{y}\right)||\pi^{\mathrm{J}}\left(\theta\right)\right]}$$

Multiple Experts

- Extending the Data Agreement Criterion (DAC)
  - Comparing multiple experts 1,…,*D*
  - Allow relative ranking

$$\mathrm{DAC_d} = \frac{KL\left[\pi^{\mathrm{J}}\left(\theta|\mathbf{y}\right)||\pi_{\mathrm{d}}\left(\theta\right)\right]}{KL\left[\pi^{\mathrm{J}}\left(\theta|\mathbf{y}\right)||\pi^{\mathrm{J}}\left(\theta\right)\right]}$$

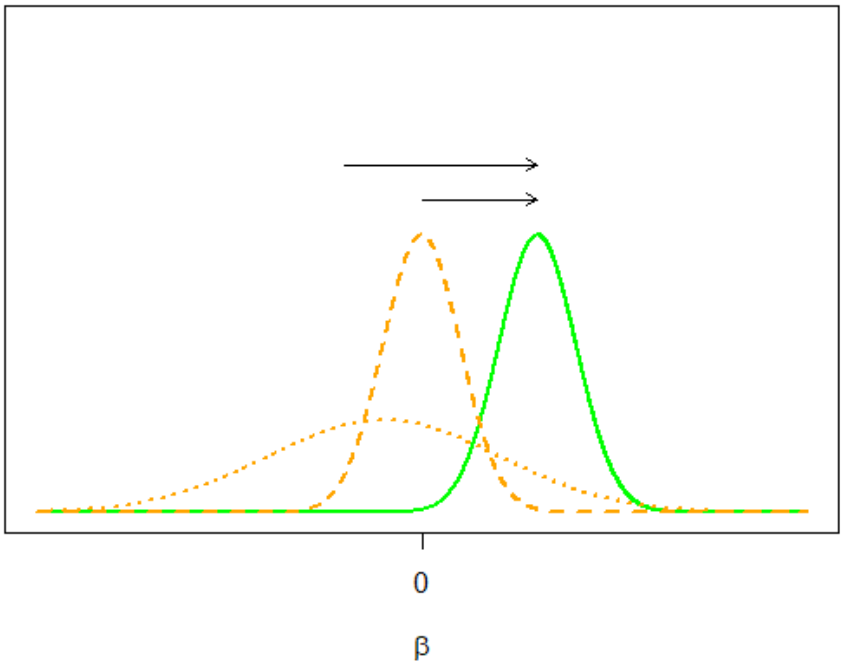
- Mathematically small step, Conceptually huge leap

Data & Expert Knowledge

- Who best predicts new data?
- Do my data and experts agree?
- Are my data reliable?

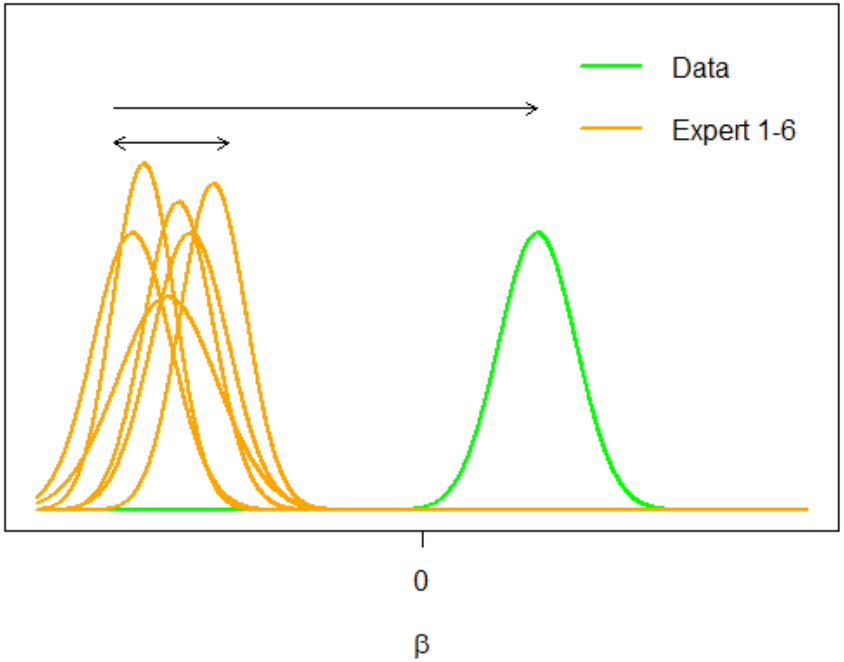
Data & Expert Knowledge

- Who best predicts new data? / Do my data and experts agree?

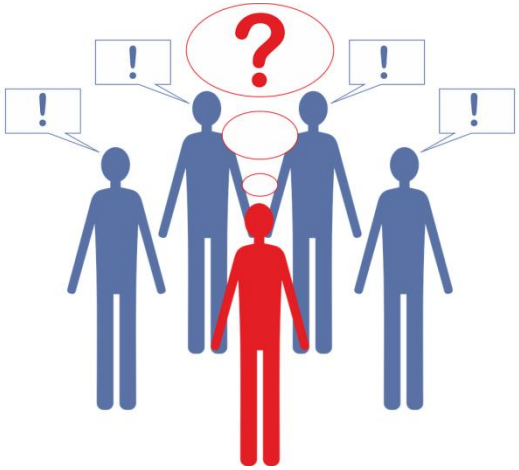


Data & Expert Knowledge

- Are my data reliable?

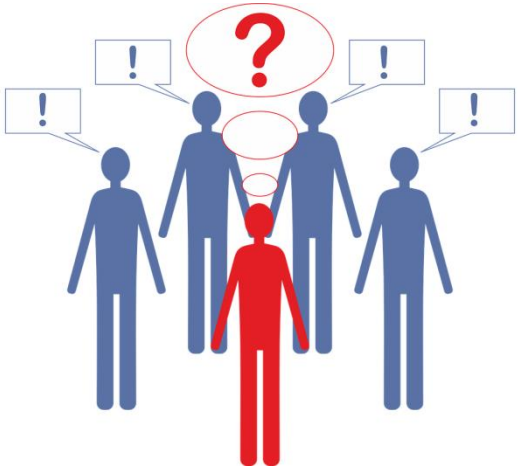


Empirical Example[4]

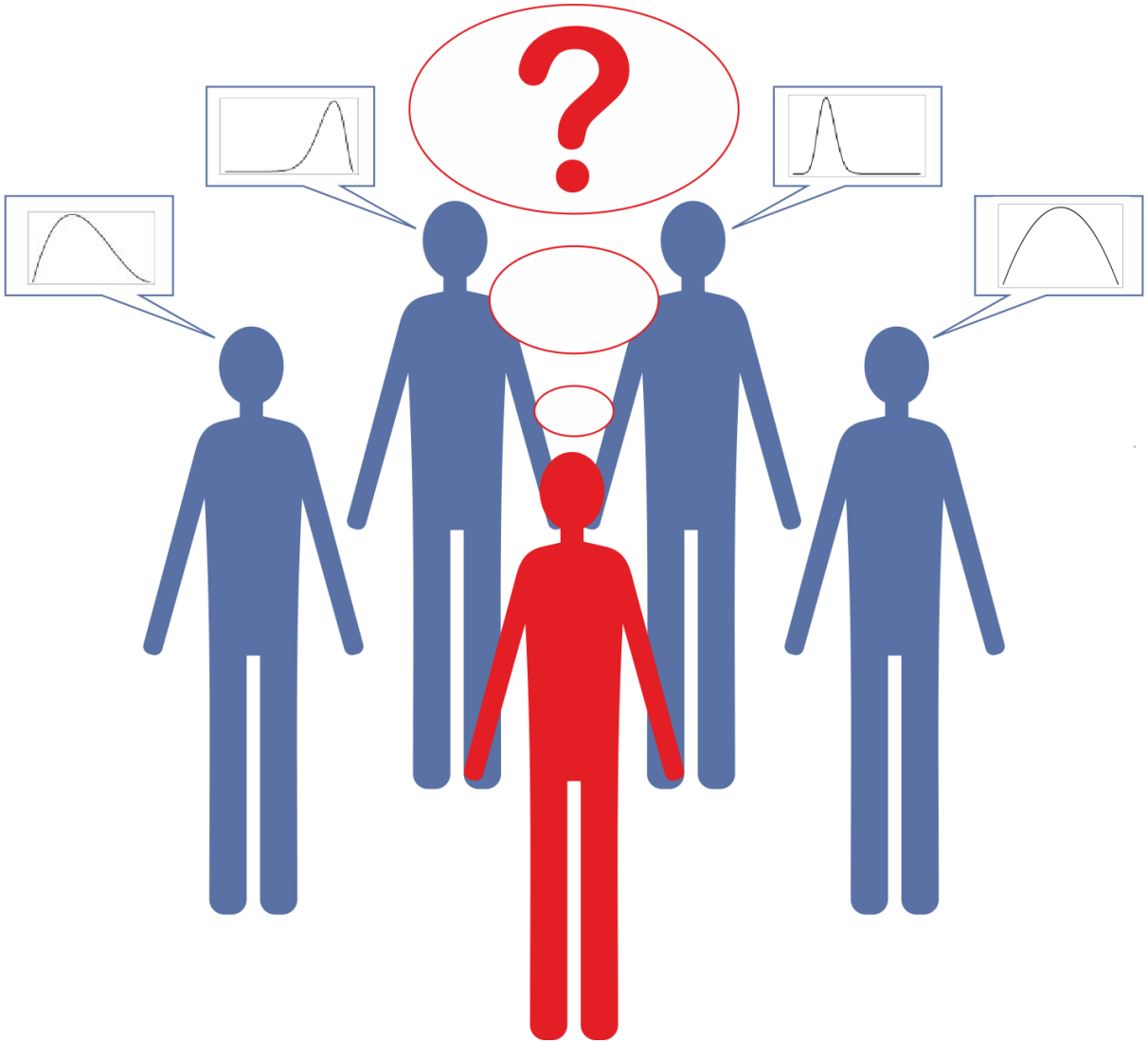


- Large financial institution
- Regional directors as experts
- Predicting average turnover per professional

Empirical Example[4]



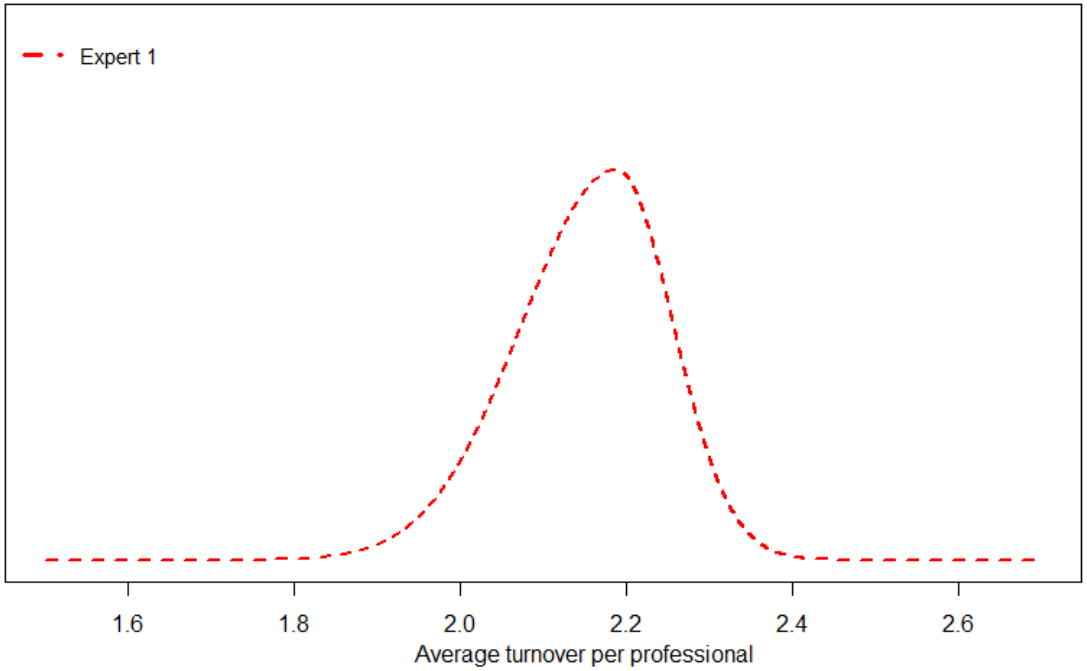
- Large financial institution
- Regional directors as experts
- Predicting average turnover per professional
- **Who best predicts new data?**



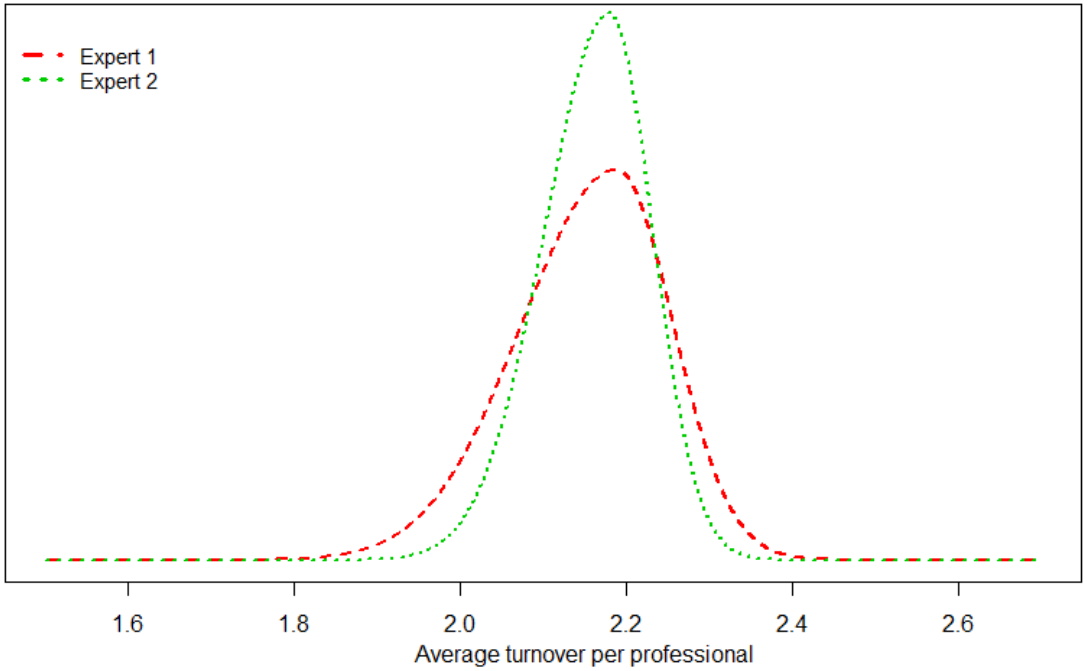
Empirical Example

- Expert Elicitation
  - Five-Step Method[5]
  - Digital elicitation environment[6]
- Transformed Data
- N(0,10000) Benchmark prior

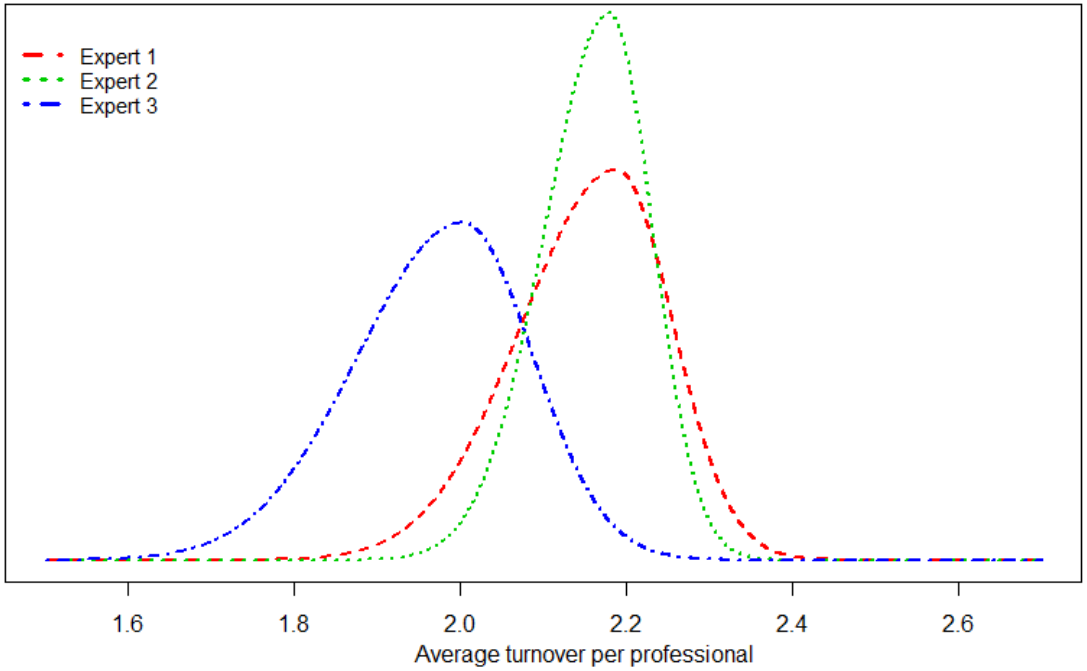
Empirical Example



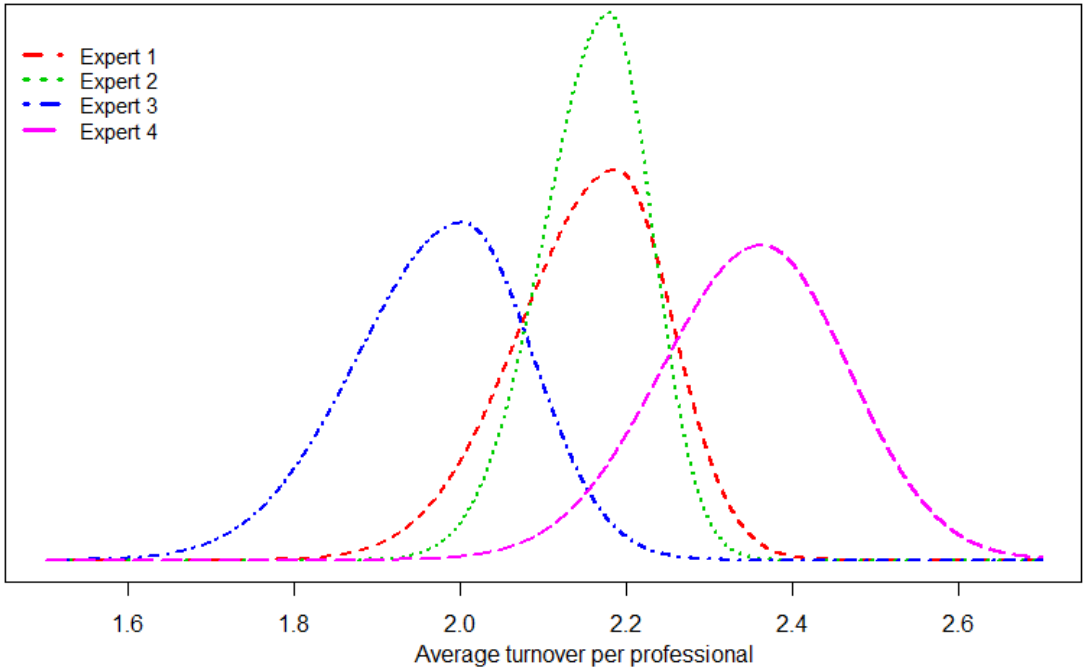
Empirical Example



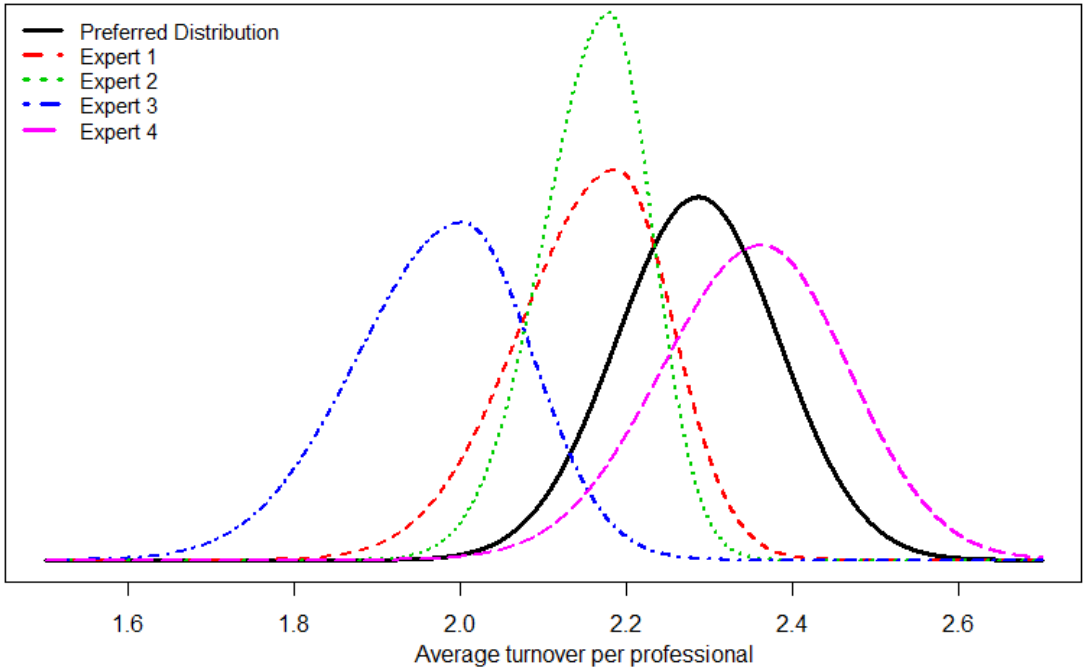
Empirical Example



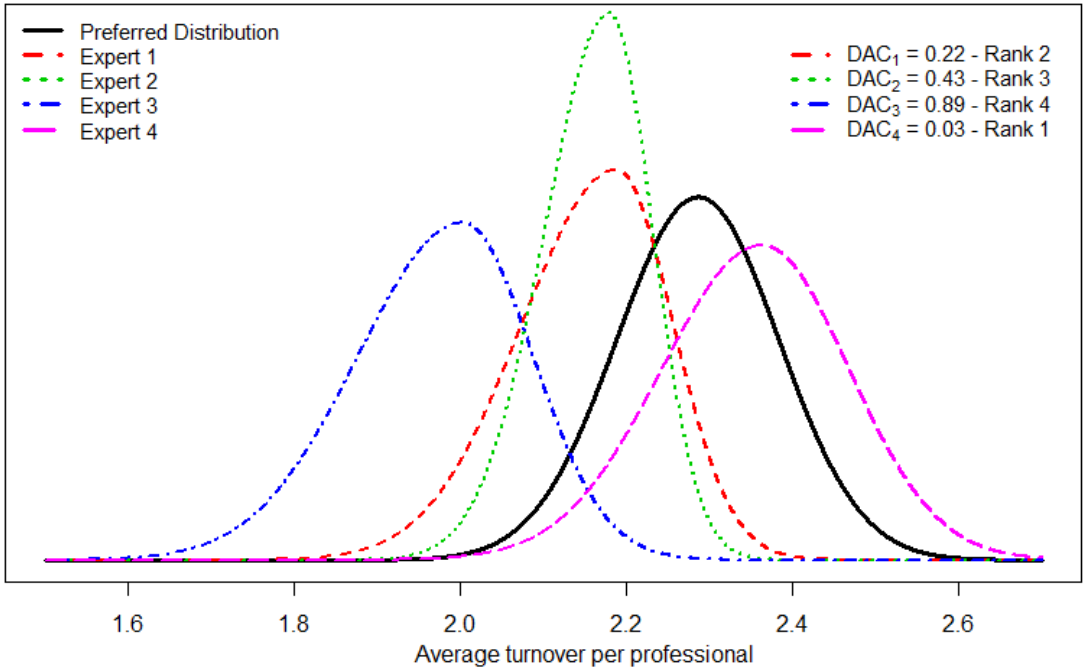
Empirical Example

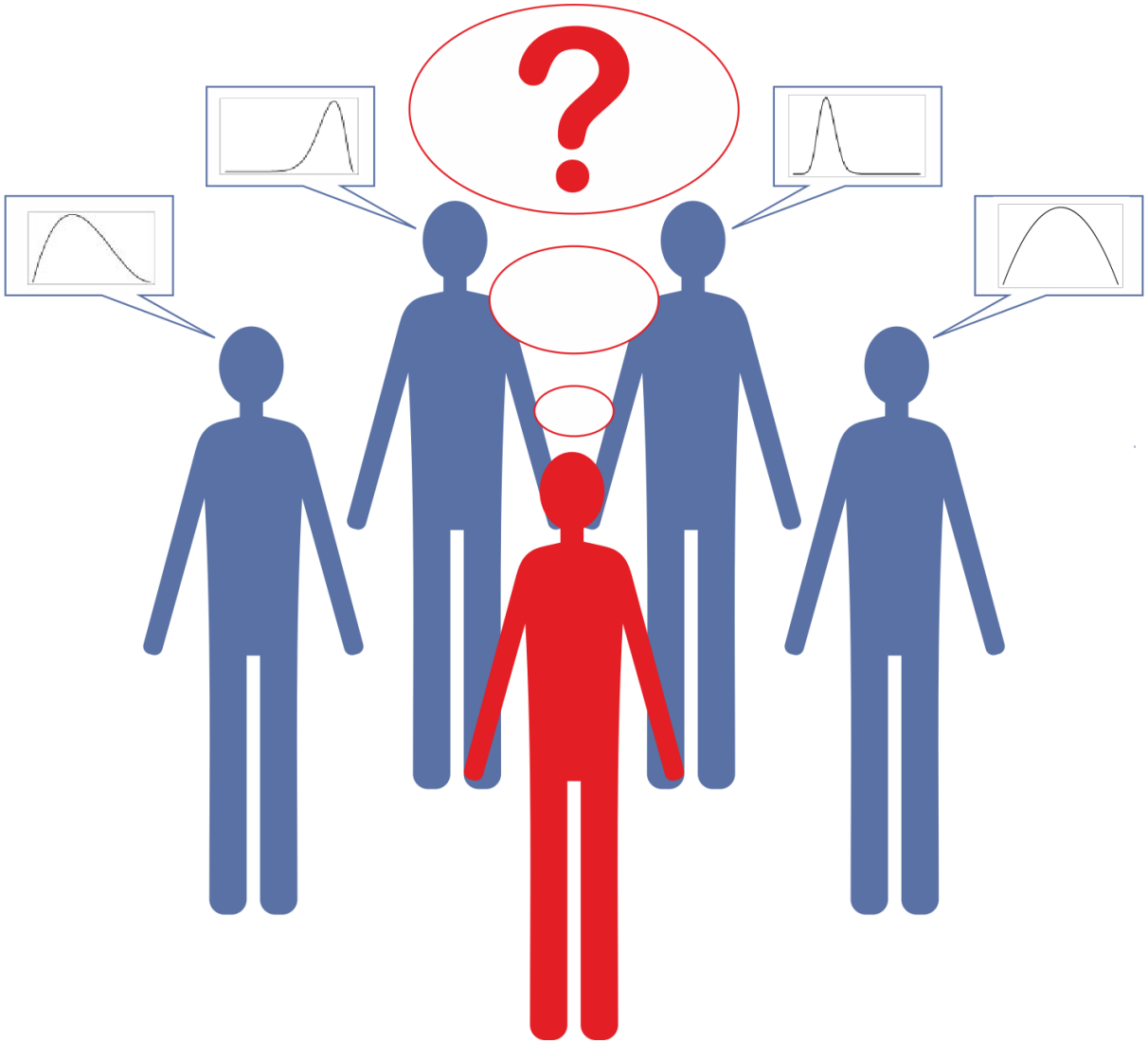


Empirical Example



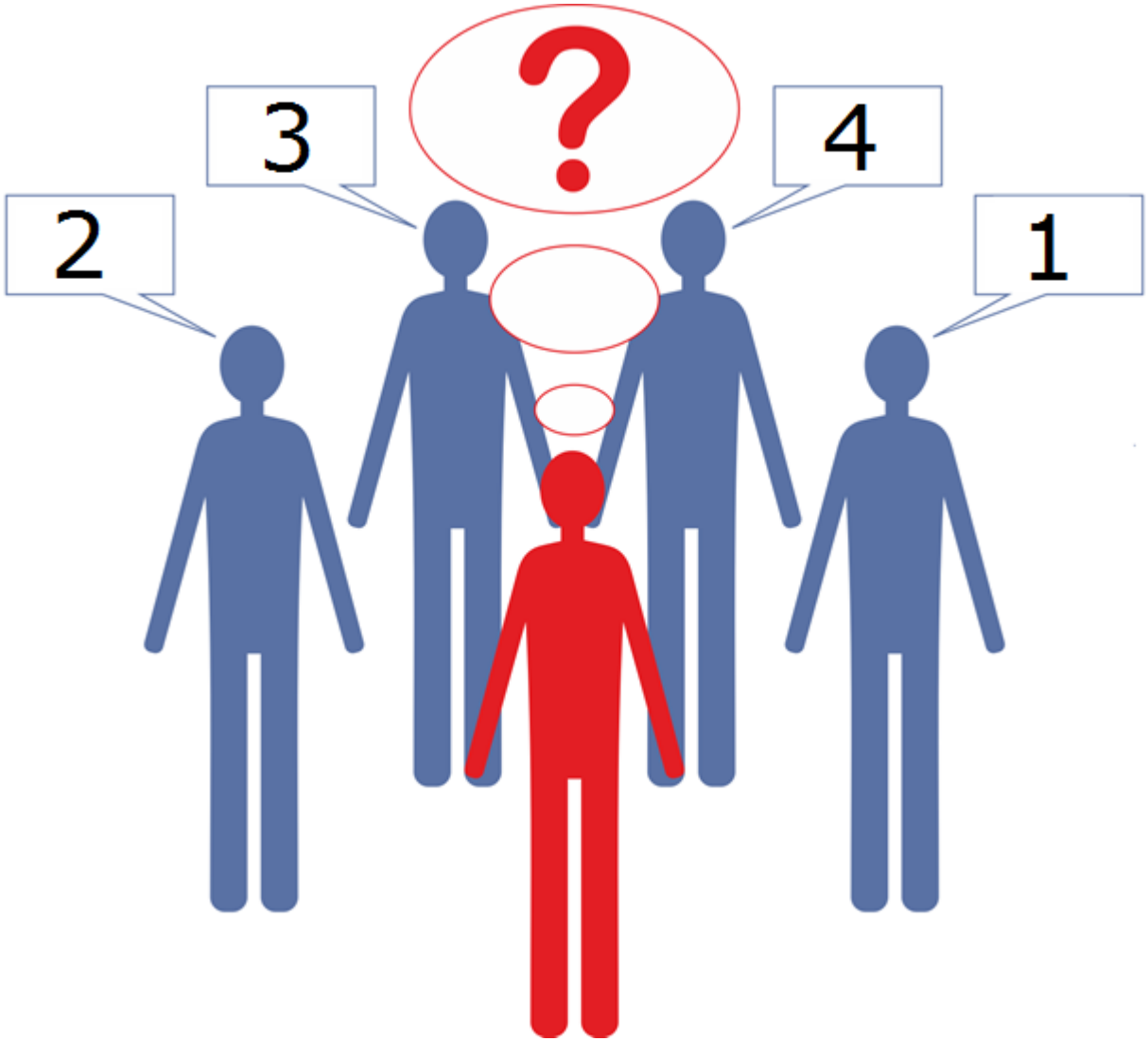
Empirical Example





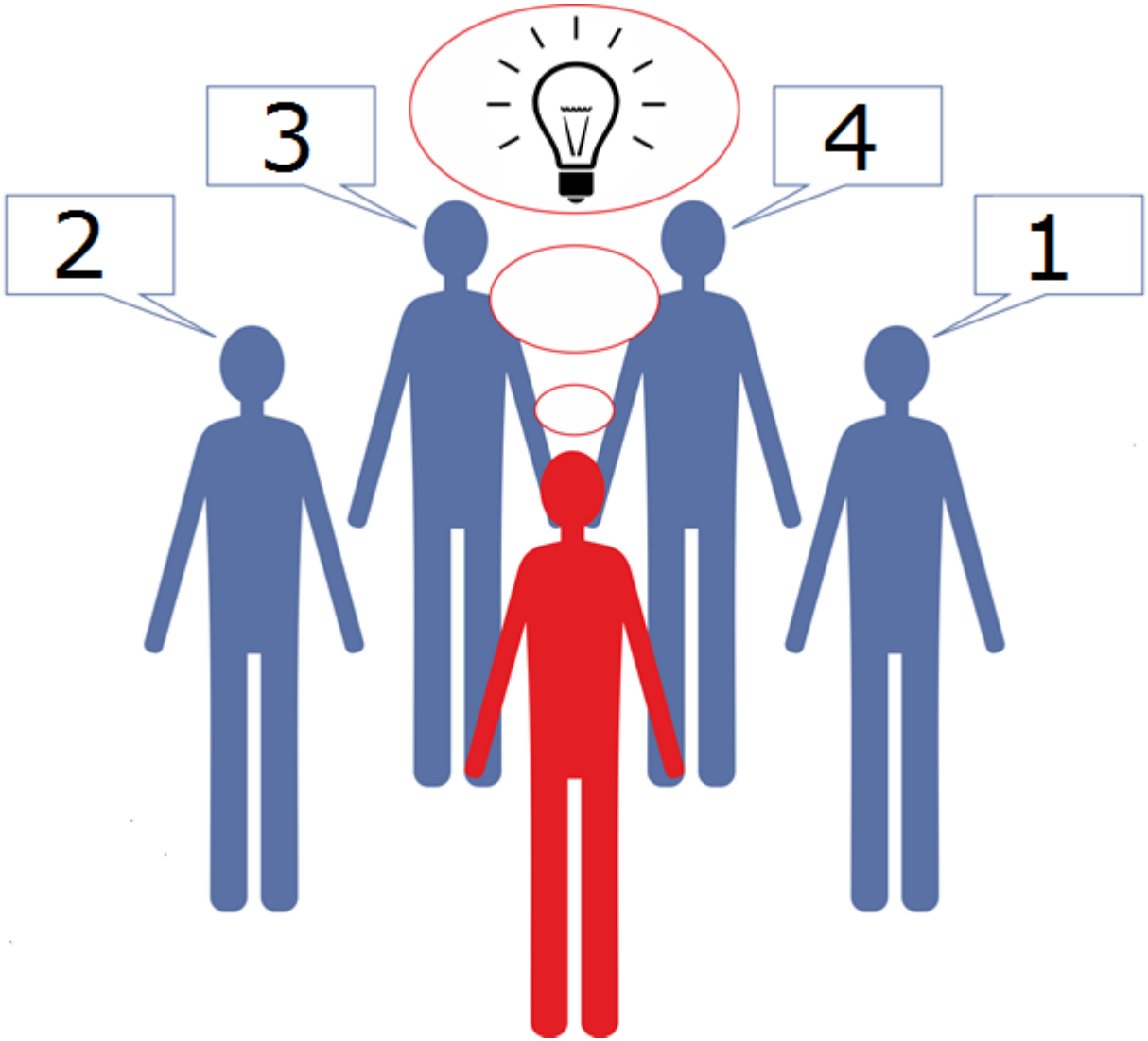
Universiteit Utrecht

profitwise



Universiteit Utrecht

profitwise



Universiteit Utrecht

profitwise

References

1. Bousquet, N. (2008). Diagnostics of prior-data agreement in applied bayesian analysis. *Journal of Applied Statistics*, 35(9), 1011–1029.
2. Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *The annals of mathematical statistics*, 22(1), 79–86.
3. Veen, D. (n.d.). KL-app. Retrieved May 17, 2017, from https://utrecht-university.shinyapps.io/klapp/
4. Veen, D., Stoel, D., & Van de Schoot, R. (2017). Using the Data Agreement Criterion to Rank Experts' Beliefs. Manuscript in preparation, Utrecht University, Utrecht.
5. Veen, D., Stoel, D., Zondervan-Zwijnenburg, M., & Van de Schoot, R. (2017). A Five-Step Method to Elicit Expert Judgement. Manuscript in preparation, Utrecht University, Utrecht.
6. Veen, D. (n.d.). Five-Step Method Elicitation App. Retrieved May 17, 2017, from https://utrecht-university.shinyapps.io/elicitation/