‘Bayesian’ and ‘Frequentist’ approaches to data-analysis

J. Hanley

Department of Epidemiology, Biostatistics and Occupational Health
McGill University

Round Table Discussion
Statistical Society of Montreal

April 23, 2010
I distinguish 3 situations

- Life in general
- Single-study data-analysis
- Research-synthesis
In Life: we use Bayes Theorem to learn/update Bayesian integration in sensorimotor learning


“When we learn a new motor skill, such as playing an approaching tennis ball, both our sensors and the task possess variability. [ ... ] We show that subjects internally represent both the statistical distribution of the task and their sensory uncertainty, combining them in a manner consistent with a performance-optimizing bayesian process. The central nervous system therefore employs probabilistic models during sensorimotor learning.”

Subconsciously, Athletes May Play Like Statisticians
What age is this person?
Additional info: He obtained his PhD 32 years earlier


Table 18 Distribution of 1997 Doctorate Recipients by Age at Doctorate

Age when earned doctorate

Age 32 years later
Single-study data-analysis

- Likelihood-based parameter-fitting + frequentist-based interval estimation
- IF intractable ML fitting issues, measurement error, hierarchical models, ..., use a computer-intensive MCMC approach using a non-informative prior.
- In simple cases, (Frequentist) confidence intervals $\approx$ (Bayesian) credible intervals. Since credible intervals involve $\text{Prob}(\theta \mid data)$ rather than $\text{Prob}(data \mid \theta)$, they are more natural, and easier to explain correctly.
- Present evidence just from study in question, preferably in a form [Likelihood] that can be merged with evidence from other studies.
Synthesis of data from several sources

- Simplest case: meta-analysis
- If no unanimity about past evidence, present range of posteriors based on range of pessimistic ↔ optimistic priors.
- Combining prior + data: is it same as adding (log) probability densities?

\[
p_{\text{photo} + \text{PhD}}(age) \propto p_{\text{photo}}(age) \times p_{\text{ageAtPhD}}(age - 32)
\]

\[
\propto p_{\text{photo}}(age) \times p_{\text{ageAtPhD}+32}(age)
\]
What age is this person?

Age, based on photo

Age, based on PhD data

Age, based on photo and PhD data
What age is this person?

![Bar chart showing age distribution based on photo.](chart1)

Age, based on photo

![Bar chart showing age distribution based on PhD data.](chart2)

Age, based on PhD data

![Bar chart showing age distribution based on both photo and PhD data.](chart3)

Age, based on photo and PhD data
What age is this person?

Age, based on photo

Age, based on PhD data

Age, based on photo and PhD data
FUNDING & CO-ORDINATES

Natural Sciences and Engineering Research Council of Canada

Le Fonds québécois de la recherche sur la nature et les technologies

James.Hanley@McGill.CA

http://www.biostat.mcgill.ca/hanley