South African Statistical Association (SASA)

56th Annual Conference

Grahamstown
October 2014

Thomas Bayes

THE MODERN DAY STATISTICS TOOLBOX

- 250th Anniversary: Essay towards solving a problem in the doctrine of chances in *Philosophical Transactions of the Royal Society of London* (1764)
- A technique: based the probability of an event to happen in a given circumstance, on a prior estimate of its probability under these circumstances

$$X \sim binomial(n;\pi) \quad \text{with} \quad p(\pi) = 1$$

- Objective Bayesian theory: Bayes, Laplace …
- Version of the Central Limit Theorem: Posterior asymptotically normal
- Many applications in physical sciences; a version of Fisher exact test

---

- Until 1838: Probability theory
- 1838 – 1950: Inverse probability (de Morgan)
- 1950 - : Bayesian analysis
- Constant prior: (Boole, Venn, …) logically unsound
- Alternatives like Fisher's developments of likelihood methods, Neyman's developments of frequentist philosophy – great appeal

- Harold Jeffreys revived objective Bayes: Jeffreys prior yielded 'accepted' procedures in all standard statistical situations; about p-values:

"A hypothesis, that may be true, may be rejected because it has not predicted observable results that have not occurred."

- 29 September 2014; Science section:

The New York Times

"The Odds, Continually Updated" by F.D. Flam

- Long Island fisherman over board in Atlantic ocean:

"owes his life to a once obscure field known as Bayesian statistics — a set of mathematical rules for using new data to continuously update beliefs or existing knowledge"

- Allowing scientists to solve problems that would have been considered impossible just 20 years ago

Do we still need to debate of being Bayesian or being frequentist?

- No agreed inferential basis for statistics - interesting for academics; but at the price of negative implications for the status of statistics in industry
- Brad Efron: 19th century – Bayesian
  20th century – frequentist
  21st century – combination Bayesian/frequentist
- Great Divide: Bayesian statistics vs frequentist / classical statistics
- Counter examples, arguments against and for both approaches
- No clear winner: Inferential views no longer compared and argued
- Bayesians debate 'objective' vs 'subjective', between siblings

- Bayesian approach: limited to small problems (high dimensional)
- Monte Carlo approaches: turned this weakness into a strength
- Three main camps: frequentists; Bayesians; pragmatists
- Asymptotic maximum likelihood inference: can be seen as a form of large-sample Bayes, with the interval for θ being interpreted as a posterior credibility interval rather than a confidence interval
- This broad view of Bayes provides a large class of practical frequentist methods with a Bayesian interpretation
- Pragmatist: good answers under both approached – two philosophies better than one (more tools for statistics toolbox)
- Unified approach – do we need this 250 year old conflicting approach?

- Example1: Test of independence

Treatment	Success	Failure
A	170	2
B	162	9

$$H_0 : \pi_A = \pi_B \quad vs \quad H_1 : \pi_A > \pi_B$$

  - Standard Pearson χ^2 - test (P)
  - Pearson test with Yates' continuity correction (Y)
  - Fisher exact test (F)
  - Bayesian solution (B) – calculates $P(\pi_A < \pi_B | data)$

$$p(\pi_A, \pi_B) \propto \pi_A^{-1/2} \pi_B^{-1/2}$$

- p-values / probability:
  - 0.016 (P)
  - 0.030 (F)
  - 0.032 (Y)
  - 0.013 (B)

- **Example 2: Single sample t test with bounded variance**

  $X \sim N(\mu, \sigma^2)$

  $X_1, \ldots, X_7$ **random sample of size** ($n = 7$): $\bar{x} = 1; s = 1$

  - **Population standard deviation (σ) unknown, the usual 95% interval for the population mean is:**

    $$I(\mu)_{95\%} : \bar{x} \pm 2.447 \frac{s}{\sqrt{7}} = 1 \pm 0.92$$

    and a Bayesian interprets it as a 95% posterior credibility interval based on Jeffreys' reference prior (RP) distribution:

    $$p(\mu, \sigma) \propto 1/\sigma$$

  - **In a measuring instrument, say $\sigma = 1.5$, the standard 95% interval:**

    $$I(\mu)_{95\%} : \bar{x} \pm 1.96 \frac{1.5}{\sqrt{7}} = 1 \pm 1.11$$

  - **Extra information ($\sigma > 1.5$), the RP modified yields 95% interval:**

    $$I(\mu)_{95\%} : \bar{x} \pm 1.45$$

- **What about unifying approaches? Something like an agreed 'hybrid approach'**
- **Emphasising calibrated Bayesian inference, George Box (1980), Donald Rubin (1984)**
- **Use the best of frequentist and Bayesian approaches**

<u>Frequentist approach</u>:

- **Separation of the role of prior information in model formulation and the role of data in estimating parameters.**
- **Treated on a more equal footing in the Bayesian approach.**
- **The frequentist approach is flexible, in the sense that full modelling is not necessarily required, and inferences lack the formal structure of Bayes' theorem under a fully specified prior and likelihood.**
- **Any method is frequentist, provided its frequentist properties can be studied.**

- It's is not prescriptive - in the sense of assessing the properties of inference and not so much about the inference system as such

- Frequentist theory is incomplete - here we can think of the interpretation with confidence intervals

- Frequentist theory is incoherent, in the sense that it may violate the likelihood principle. The likelihood principle plays an important role in the inferential debate since it is satisfied by Bayesian inference and violated by frequentist inference

- Example: Bernoulli(π) experiment
  - Binomial sampling, repeated n=12 times and X=3 successes observed
  - Negative binomial sampling, repeated until x=3 successes on N=12 repetitions
  - Likelihood:

$$L(\pi) \propto \pi^3 (1-\pi)^9$$

  - Same inference? (same MLE), but

$$H_0 : \pi = 0.5 \quad vs \quad H_1 : \pi < 0.5$$

Sampling space different , different p-values (0.073 vs 0.033)

Bayesian approach:

- For a given Bayesian model and prior distribution, Bayes' theorem is the simple prescription that supplies the inference
- It may be difficult to compute, and checks are needed to ensure that the posterior distribution is proper, but the solution is clear
- Bayes is viewed as too subjective for scientific inference
- methods include methods based on non-informative priors that some classify as frequentist
- Bayes requires and relies on full specification of a model (likelihood and prior)
- Bayes yields *too many answers*

- Crudely and over-simplistically:
  - Bayesian statistics is strong for inference under an assumed model, but is relatively weak for the development and assessment of models
  - Frequentist statistics provides a useful tool for model development and assessment, but is a weak tool for inference under an assumed model
  - So, a natural compromise is to use frequentist methods for model development and assessment, and Bayesian methods for inference under a model

- *Example 1 (continued).*
  - The conjugate Bayesian inference adds Beta priors for the success rates in the two groups and computes $P(\pi_A > \pi_B|data)$
  - Proper prior distributions may be considered in certain contexts; when there is little prior evidence about the success proportions, the choice of '*objective prior*' has been debated, but Jeffreys' prior is one plausible conventional choice
  - This hybrid Bayesian method limits the vagueness in the reference set for frequentist assessments to model evaluation, rather than to model inference under a specified model.

- *Example 2 (continued):*
  - Posterior predictive distribution of the sample variance S^2: The posterior probability of observing a sample variance in future datasets as low as the observed sample variance of $s^2 = 1$ is about 0.065, which is low but not exceptional, so the Bayesian model seems not unreasonable
- This is not to claim that this hybrid Bayesian method solves all the problems of statistical inference
- Ambiguities will arise when combining model inference and model checking
- How much peeking at the data is allowed in developing the model without seriously corrupting the inference?

<u>What are the implications for teaching and in practice of statistics?</u>

- **More emphasis should be on statistical modelling than on statistical methods.**
- **Models with better fits can yield worse predictions than methods that fit the observed data better**
- **More attention is needed to assessments of model fit. This is where frequentist methods have an important role**

<u>What are the implications for teaching and in practice of statistics?</u>

- ***Bayesian statistical methods need to be taught.***
  - **Bayesian statistics is absent or "optional" in many programs for training masters and even Ph.D. statisticians**
  - **A Bayes course should be a required component of any masters and Ph.D. programme in statistics**
  - **If we think of the consumers of statistics (the service courses), Bayes is not a part of these services courses, so most think of frequentist statistics as the only approach, and are not aware that Bayesian inference exists**
  - **The basic idea of Bayes' theorem does not require calculus, and Bayesian methods can be teach if the emphasis is placed on interpretation of models and results, rather than on the inner workings.**

What are the implications for teaching and in practice of statistics?

- Bayesian posterior credibility intervals have a much more direct interpretation than confidence intervals – our consumers of statistics in any case interpret it this way to their managers
- Any case, frequentist hypothesis testing is no picnic to teach to these consumers of statistics.

Bayesian and frequentist ideas are important for good statistical inference, and both sets of ideas need to be developed and taught.