Elliott Sober on Coincidence

Wouter Meijs
Erasmus University Rotterdam
The Netherlands

Akaike's Independence Criterion

AIC appears to present a promising approach to the question of coincidence, which is both objective and constructive

Outline

1. Definition of Coincidence
2. Deriving AIC
3. Some possible defenses for Bayesianism

Possible definitions of Coincidence

- Diaconis and Mosteller:
  - Surprising
  - Perceived as related
  - No apparent causal connection
- Sober:
  - No common cause

Birthday Problem

Two situations

1. Four persons in this room have the same birthday
2. No persons in this room have the same birthday

Conclusion

- Conceptual analysis of the common sense notion of coincidence reveals subjective aspects
- Sober's analysis: what is a scientifically interesting notion of coincidence?

Continental Drift Example

1. Data:
    a. Similarities in coastline
    b. Genetic similarities
    c. Similarities in geological strata
2. Two models:
    i. Coincidence
    ii. Continental Drift

Deriving AIC

- $$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right]$$
- Equal variance σ^2 ?
- Measurement errors?

Some Possible Defenses for Bayesianism

1. Comparative judgments
2. Examples
3. Simplicity

Comparative Judgments I

- $H =_{def}$ The result is not a coincidence
- $H_0 =_{def}$ The result is a coincidence
- Strategy: prefer H over H_0 iff:

$$p(H \mid E) > p(H_0 \mid E)$$

Comparative Judgments II

- $$\frac{p(H \mid E)}{p(H_0 \mid E)} = \frac{p(E \mid H)}{p(E \mid H_0)} \cdot \frac{p(H)}{p(H_0)}$$
- $$p(H \mid E) > p(H_0 \mid E) \iff \frac{p(E \mid H)}{p(E \mid H_0)} > \frac{p(H_0)}{p(H)}$$
- $$\frac{p(H_0)}{p(H)} \iff \frac{p(E \mid H)}{p(E \mid H_0)}$$

The New Jersey Lottery I

- H: The NJL was rigged.
- H_0: The NJL was fair.
- E: In 2003 somebody won the NJL twice within 4 months.

The New Jersey Lottery II

- Samuels/McCabe: odds are 1 in 30
- $$\frac{p(E \mid H)}{p(E \mid H_0)} < \frac{1}{0.033}$$
- $$p(H \mid E) > p(H_0 \mid E) \iff p(H_0) < 30\, p(H)$$

Continental Drift I

- H_1: Africa and South-America were once connected.
- H_0: Africa and South-America have never been connected.
- E_1: Genetic similarities
- E_2: Geological similarities
- E_3: Coastline similarities.

Continental Drift II

- $$p(E \mid H_0) = p(E_1 \mid H_0)\, p(E_2 \mid H_0)\, p(E_3 \mid H_0) < \left(\frac{1}{1000}\right)^3$$
- $$p(H \mid E) > p(H_0 \mid E) \iff \frac{p(H_0)}{p(H)} > 10^9$$

Simplicity

1. Many parameters
2. All parameters must be taken into account
3. Difficult calculations

Conclusion: AIC and Coincidence

AIC appears to present a promising approach to the question of coincidence, which is both objective and constructive.

However, it is a complicated approach and it is not yet clear that it performs better than Bayesianism.