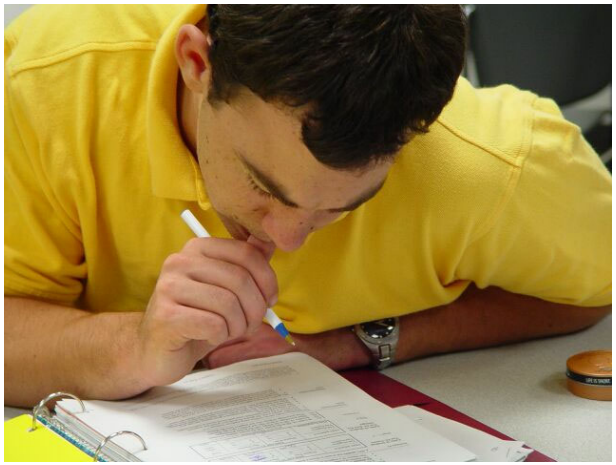
Statistical Literacy©

critical thinking about arguments involving statistics as evidence

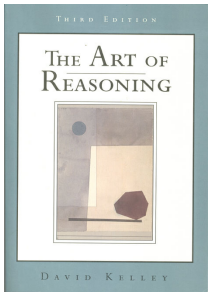


PROJECT OVERVIEW

StatLit Project Brochure1k 6/12/2003 ©

Statistical Literacy
The Arguments

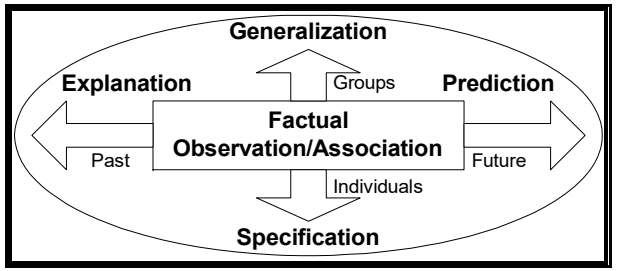
Statistical Literacy is the ability to read and evaluate statistically-based arguments found in everyday life.



Statistical Literacy is critical thinking about statistics as evidence in arguments. In observational studies association is contextual. Verbally the difference can be obscure. Consider three statements:

A. Subjects with more X have more(less) Y.
B. As X increases, Y increases (decreases).
C. Increasing a subject's X increases (decreases) their Y.

A is association; C is causation. Technically, B is association; non-technically it's causal.



In moving from association to causation, three distinctions are considered primary:

- Material (manipulative) vs. formal (shift in focus) in explanation and prediction,
- Experiment versus Observational Study,
- Sample versus Population.

Statistical Literacy:
The Evidence

Assuming a statistic is appropriate, the strength it provides as evidence in arguments depends on the influence of **chance** (randomness), of **bias** (systematic error) and of **confounding** (relevant factors not taken into account).

- The strength of statistics based on *small-sized, well-designed* ***experiments*** (treatments with random assignment) is most likely to be influenced by *chance*.
- The strength of statistics based on *poorly-designed studies* is most likely to be influenced by *bias* (systematic error).
- The strength of statistics based on *large-scale, well-designed* ***observational studies*** *(non-experiments)* is most likely to be influenced by *confounding*.

Observational studies are much more common than *experiments* in business, in social policy issues, and in the popular press. *Large-scale, well-designed observational studies* are becoming increasing common.

Conclusion

To be statistically literate about statistics as evidence in everyday arguments, students must take CARE to be aware of *Confounding, Appropriateness, Randomness (chance), and Error (bias).*

Statistical Literacy:
Seeing Confounding

Statistical Literacy focuses on "*seeing the story behind the numbers*": thinking about what factors may be ***confounding*** (confusing) the numbers. This means treating associations as contextual.

- Count Compare: A judicial system may be biased against non-whites even though there are more whites in prison (given more whites in the population).
- Ratio Compare: A judicial system may be fair even if non-whites are more likely to be in prison than whites (non-whites may be more likely to be poor than whites and thus more likely to be in prison).
- Ratio Compare: The hospital with the highest rate of deaths may be the best (more patients may be in poor condition).
- Linear Model: For each additional bedroom, the price of houses increases by $39,000 (p=0.00). After taking into account land values, square footage and number of baths, the price of houses increases by $5,000 per bedroom (p=0.10).

Moral

Including an additional confounder can change the direction or the amount of an arithmetic association in **any** observational study. Since these associations are contextual, one must always check their context!

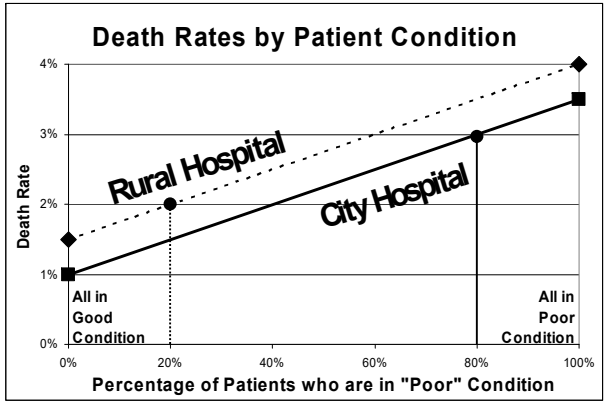
Statistical Literacy:
Controlling Confounding

Statistical Literacy focuses on quantitative techniques (ratios, comparisons and **standardization**) to take into account – to control for – the influence of *confounding.*

Using ratios and comparisons includes:

- Describing rates and percentages
- Decoding tables of rates & percentages
- Comparing rates and percentages

Standardization example: The death rate in a city hospital was *higher* than a rural hospital (3% vs. 2%). But for patients in good condition, the death rate in the city hospital was *lower* than in the rural hospital (1% vs. 1.5%). The same was true for patients in poor condition (3.5% vs. 4.0%). *So, the standardized death rate* in the city hospital was actually *lower* than the rural hospital.



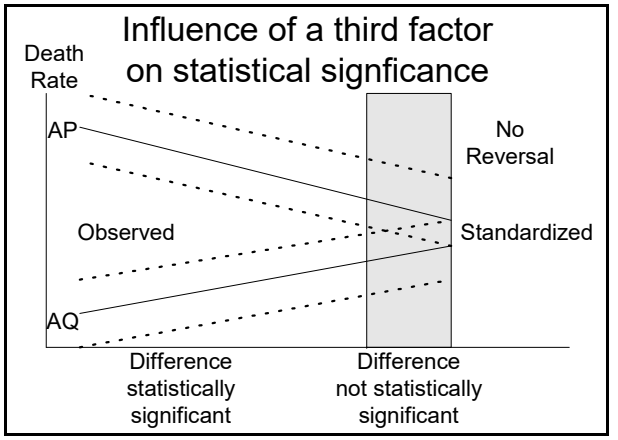
Standardization techniques include standardized scores, standard age populations, multivariate regression and analysis of covariance.

Statistical Literacy
Statistical Significance

To show that statistical significance is contextual in observational studies, a simplified approach using confidence intervals for proportions is used. See Chapter 5, *Understanding Scientific Reasoning* by Giere.

A difference that is statistically significant is indicated by a lack of overlap for the associated confidence intervals.

The confidence intervals used are conservative so they depend only on the confidence level (Z) and the size of the sample (n).



A statistically significant association can become statistically significant after taking into account a confounder's influence – and vice versa.