

# LIKELIHOOD-BASED STATISTICAL DECISIONS

Marco Cattaneo  
Seminar for Statistics  
ETH Zürich, Switzerland

July 23, 2005

Let  $\mathcal{P}$  be a **set of statistical models**:

- $\mathcal{P}$  is a set of probability measures on a measurable space  $(\Omega, \mathcal{A})$ ;
- absolutely no structure is imposed on  $\mathcal{P}$  (for instance,  $\mathcal{P}$  could be a nonparametric family of models).

The **likelihood function**  $lik_A : \mathcal{P} \rightarrow [0, 1]$  based on the observation  $A \in \mathcal{A}$ :

- is defined by  $lik_A(P) = P(A)$ ;
- measures the relative plausibility of the models  $P \in \mathcal{P}$ , on the basis of the observation  $A$  alone;
- is not calibrated: only ratios  $lik_A(P)/lik_A(P')$  are well determined in a statistical sense.

# Relative Plausibility

The **relative plausibility**  $rp$  on  $\mathcal{P}$  generated by  $lik_A$ :

- is the class of nonnegative functions on  $2^{\mathcal{P}}$  defined by  $rp(\mathcal{H}) \propto \sup_{P \in \mathcal{H}} lik_A(P)$ ;
- is a non-calibrated possibility measure on  $\mathcal{P}$ .

The **description of the uncertain knowledge about the models** by means of relative plausibility:

- can be easily updated (since  $lik_{A \cap B} = lik_A lik_{B|A}$ );
- allows a natural incorporation of prior information: independent pieces of information can be combined, and complete (or partial) ignorance can be described;
- is parametrization invariant;
- satisfies the strong likelihood principle; but can also use pseudo likelihood functions;
- can be used for inference (maximum likelihood estimator, tests and confidence regions based on the likelihood ratio statistic, . . . ) and decision making (MPL criterion);
- leads to conclusions which are in general weaker than those based on a probabilistic description of the uncertain knowledge about the models; but is based on weaker assumptions, is simpler and more intuitive.

# MPL Criterion

A **statistical decision problem** is described by a loss function  $L : \mathcal{P} \times \mathcal{D} \rightarrow [0, \infty)$ :

- $\mathcal{D}$  is the set of possible decisions, and  $\mathcal{P}$  is the set of considered statistical models;
- $L(P, d)$  is the loss we would incur, according to the model  $P$ , by making the decision  $d$ .

If the uncertain knowledge about the models is described by the relative plausibility  $rp$  on  $\mathcal{P}$ , the **MPL criterion** for choosing a decision  $d \in \mathcal{D}$  consists in minimizing  $\sup_{P \in \mathcal{P}} rp\{P\} L(P, d)$ :

- the minimized quantity is the Shilkret integral of  $L(\cdot, d)$  with respect to  $rp$ : this is intuitive and simple (allowing decisions even in difficult problems);
- if invariance with respect to translations of the loss function is needed, the integral of Choquet should be used instead of the one of Shilkret;
- the obtained decision functions are equivariant (if the problem is invariant) and asymptotic optimal (if some regularity conditions are satisfied);
- the consideration of many examples suggests that the MPL criterion leads in general to reasonable decisions.

## Example

Estimation of the variance components in the  $3 \times 3$  random effect one-way layout, under normality assumptions and weighted squared error loss.

$$X_{ij} = \mu + \alpha_i + \varepsilon_{ij} \quad \forall i, j \in \{1, 2, 3\}$$

Normality assumptions:

$$\alpha_i \sim \mathcal{N}(0, v_a), \quad \varepsilon_{ij} \sim \mathcal{N}(0, v_e), \quad \text{all independent}$$

$$\Rightarrow X_{ij} \sim \mathcal{N}(\mu, v_a + v_e) \text{ dependent, } \mu \in \mathbb{R}, \quad v_a, v_e \in \mathbb{R}^+$$

The estimates  $\widehat{v}_e$  and  $\widehat{v}_a$  of the variance components  $v_e$  and  $v_a$  are functions of

$$SS_e = \sum_{i=1}^3 \sum_{j=1}^3 (x_{ij} - \bar{x}_{i.})^2 \quad \text{and} \quad SS_a = 3 \sum_{i=1}^3 (\bar{x}_{i.} - \bar{x}_{..})^2,$$

where

$$\bar{x}_{i.} = \frac{1}{3} \sum_{j=1}^3 x_{ij}, \quad \bar{x}_{..} = \frac{1}{9} \sum_{i=1}^3 \sum_{j=1}^3 x_{ij},$$

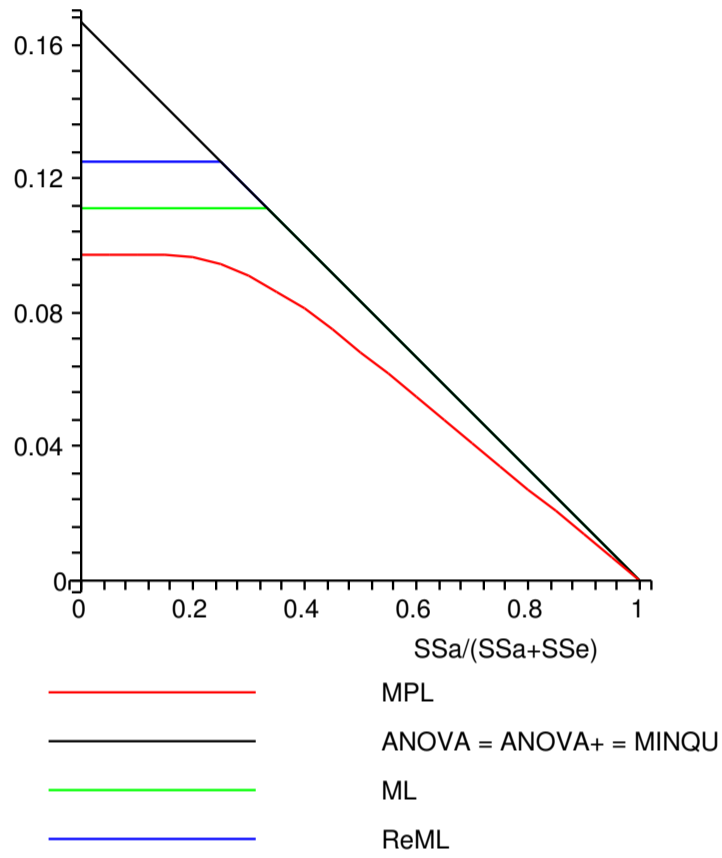
$$\frac{SS_e}{v_e} \sim \chi_6^2 \quad \text{and} \quad \frac{\frac{1}{3} SS_a}{v_a + \frac{1}{3} v_e} \sim \chi_2^2.$$

The considered loss functions are

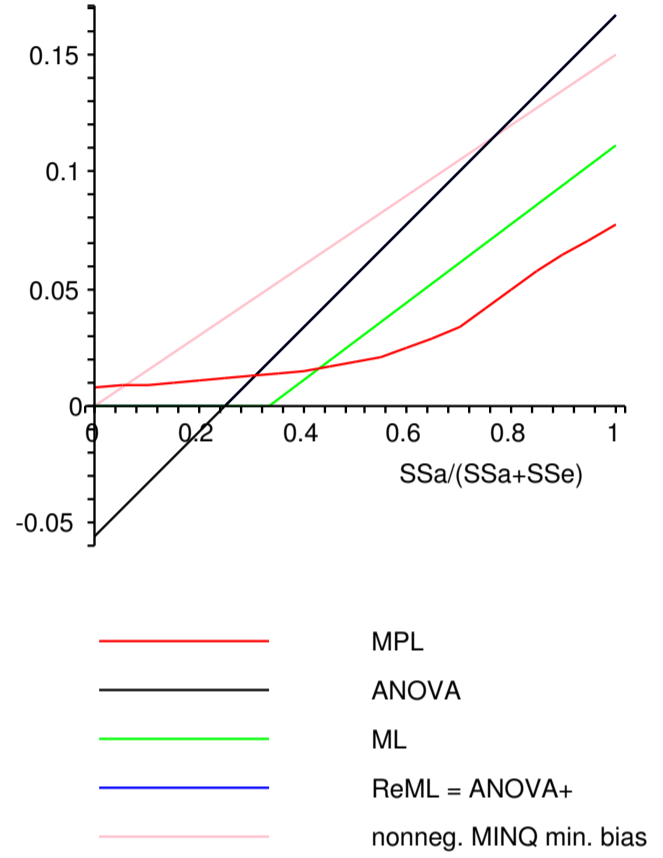
$$3 \frac{(\widehat{v}_e - v_e)^2}{v_e^2} \quad \text{and} \quad \frac{(\widehat{v}_a - v_a)^2}{(v_a + \frac{1}{3} v_e)^2}.$$

# Example

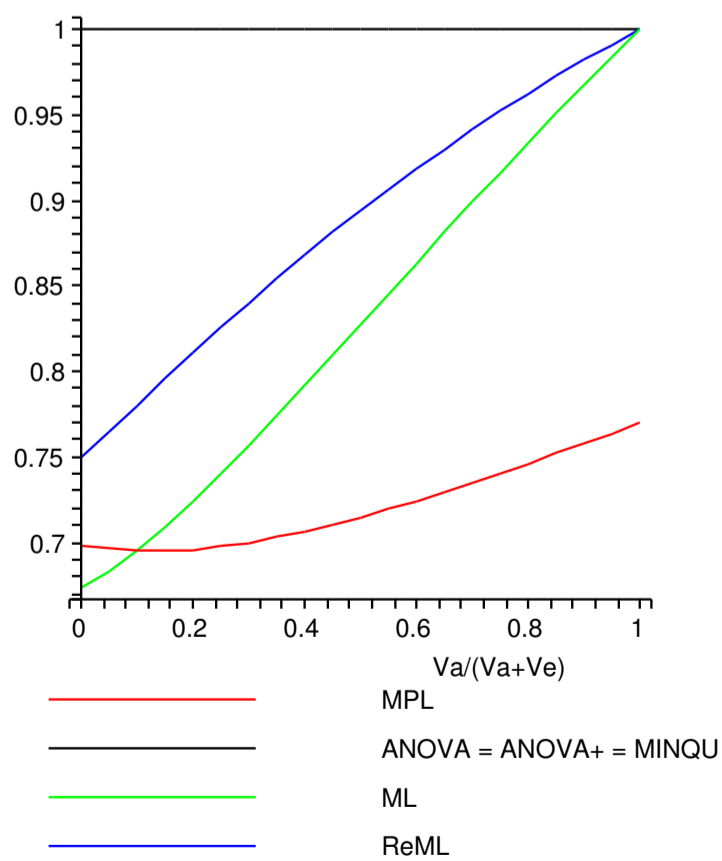
$$\frac{\widehat{v}_e}{(SS_a + SS_e)}$$



$$\frac{\widehat{v}_a}{(SS_a + SS_e)}$$



$$3 \frac{E[(\widehat{v}_e - v_e)^2]}{v_e^2}$$



$$\frac{E[(\widehat{v}_a - v_a)^2]}{(v_a + \frac{1}{3}v_e)^2}$$

