STATISTICAL METHODS FOR ANALYSING PREFERENCES IN FOREST PLANNING

Pekka Leskinen
Finnish Forest Research Institute
University of Joensuu, Department of Statistics

Saariselkä 19.4.2001

CONTENTS

1. INTRODUCTION

Consider relative merits of attributes $1, 2, ..., n$.

For example, alternative forest plans

1: Continue natural growth with no cuttings,

2: "Normal" forestry guidelines,

...

n: Maximise timber harvesting income.

Which forest plan please the forest owner most? How much better it is compared to others?

Assume that forest owner (decision maker) is interested only in scenic beauty of forest landscape.

On way to measure the preferences: Forest owner compares landscapes (computer images) $1, 2, ..., n$ **pairwisely**.

There are $m = n(n-1)/2$ pairs. Decision maker gives the relative merit for each pair (or some of the pairs) in a verbal scale

"equal", "weak", "strong", "very strong", absolute".

Numerical counterparts for verbal expressions?

Several alternatives to numerical measurement scale have been proposed.

	equal	weak	strong	very s.	absol.
S:	1/1	3/1	5/1	7/1	9/1
L:	1/1	1.7/1	3/1	5.2/1	9/1
S & H:	1/1	1.5/1	2.3/1	4/1	9/1
M & Z:	1/1	1.3/1	1.8/1	3/1	9/1

How statistical inference depends on the choice of the measurement scale?

2. REGRESSION MODEL

Let r_{ij} be the relative value of attribute i compared to attribute j given by the judge in some numerical measurement scale.

Assume that $r_{ij} = (v_i / v_j)\exp(e_{ij})$, where v_i and v_j are the true values of attributes i and j, and e_{ij} measures the uncertainty or error.

Defining $y_{ij} = \log(r_{ij})$, the regression model for the pairwise comparisons data gets the form

$$y_{ij} = a_i - a_j + e_{ij}, \qquad (1)$$

where $a_i = \log(v_i)$, and the residuals are uncorrelated with $\mathrm{E}(e_{ij}) = 0$, and $\mathrm{Var}(e_{ij}) = s^2$. The model can be written as $\mathbf{Y} = \mathbf{X}\mathrm{a} + \mathrm{e}$, so $\hat{\mathrm{a}} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{Y}$, and $\hat{s}^2 = (\mathbf{Y} - \mathbf{X}\hat{\mathrm{a}})^{\mathrm{T}}(\mathbf{Y} - \mathbf{X}\hat{\mathrm{a}})/(m - n + 1)$.

Extensions of Eq. (1):
- Multiple judges.
- Multiple decision criteria.
- Interval judgments.
- Background characteristics of judges or attributes.

3. MEASUREMENT SCALES AND CONSISTENCY

If pairwise comparisons are consistent, i.e. $r_{ij} = r_{it} r_{tj}$ for every i, j and t, then $\hat{s}^2 = 0$. Otherwise $\hat{s}^2 > 0$.

Inconsistency is natural feature, because it is difficult to express preferences.

Consider measurement scale with n steps and attributes $i = 1, 2, ..., n$ so that $1 = v_1 < v_2 < ... < v_n$. It can be shown that only geometric measurement scale $v_i = \exp[(i-1)s]$, $s > 0$, enables consistency.

In geometric measurement scale, $\hat{s}^2 > 0$ can be interpreted as the uncertainty of the judge, because consistency would have been possible.

In other measurement scales, $\hat{s}^2 > 0$ can be also caused by the measurement scale itself.

4. INCORRECT CHOICE OF MEASUREMENT SCALE

Assume that one of the four m.scales is correct, but we don't know which one. What can be done?
(1) If the measurement scales don't differ in statistical inference, the choice doesn't matter.
(2) Use all measurement scales simultaneously to have "safe" statistical inference.
(3) Choose measurement scale, that minimise the effect of incorrect choise.

Table 1. $n=3$, *t*-test, $\mathrm{H}_0 : a_1 = a_2$, $\mathrm{H}_1 : a_1 \neq a_2$, critical value 0.05. Probabilities of some events.

Cons.	H_0	H_1	H_0, or H_1
0.08	0.85	0.00	0.07

Table 2. $n=3$, *t*-test, $\mathrm{H}_0 : a_1 = a_2$, $\mathrm{H}_1 : a_1 \neq a_2$. Probabilities that different measurement scales produce smallest (largest) *p*-value.

Cons.	S	L	S & H	M & Z	No
0.08 (0.08)	0.42 (0.23)	0.06 (0.08)	0.03 (0.03)	0.21 (0.36)	0.22 (0.22)

(1) Measurement scales are different. **(2)** Problematic.
(3) L and S & H are safer than S and M & Z.

5. EMPIRICAL COMPARISON OF MEASUREMENT SCALES

Organise an experiment, where the true values v_i / v_j are known. Areas of figures, for example.

Let $d_{ij} = 0, ..., 8$ be an index for verbal judgments (0: equal, ..., 8: absolute).

In geometric measurement scale, use model $\log(v_i / v_j) = s\, d_{ij} + e_{ij}$, and estimate optimal scale parameter $\hat{s}$.

Correspondingly, in arithmetic measurement scale, for example, $\log(v_i / v_j) = \log(s'\, d_{ij} + 1) + e'_{ij}$, and $\hat{s}'$.

Basic result: Accuracy of the estimated preferences is essentially a question of optimal scaling only.

However, the functional form of measurement scales is important in some other contexts as shown earlier.