Arbeidsnotater

S T A T I S T I S K   S E N T R A L B Y R Å

WORKING PAPERS FROM THE CENTRAL BUREAU OF STATISTICS OF NORWAY

IO 69/9 Oslo, 27 May 1969

ON THE EFFICIENCY OF GENERAL TABLEPROGRAMMES.

by

Ib Thomsen

C O N T E N T S

1. INTRODUCTION

The purpose of this paper is to study the effect of the development and use of general tabelprogrammes in the production of tables.

At the Central Bureau of Statistics of Norway has lately been developed some new general tableprogrammes, and in this paper we have tried to describe some of our experiences by a model, which might be of some interest also for other organizations.

Our first problem was to decide whether we should use one of the existing general tableprogrammes or develope a new one adjusted to our table structure. We had to find a measure for the efficiency of several existing general tableprogrammes. Section 4 presents a study on this problem. If we decided to develope our own programme, a second problem would be to decide how general a programme we should aim at. From experience we knew that it is more efficient to make a general programme which does only generate a part of our tables, and leave the rest for special programming. Some of our ideas on this problem is also incorporated in the model below.

2. THE EFFECT OF THE GENERAL TABLEPROGRAMME ON THE PROGRAMMINGTIME

Introduce

- p: The fractional part of the total number of tables generated by the general tableprogramme. $0 \leq p \leq 1$
- t: The number of hours used to develope and programme the general tableprogramme.
- L: The number of months we expect to use the general tableprogramme before it will be obsolete.
- T: The number of hours used pr. month to programme the tables if we do not use the general tableprogramme.

The number of programminghours used during the periode of L months applying partly the general programme, can be written

$$O_1 = t + (1 - p)\, L \cdot T$$

Here $(1 - p)\, LT$ is the time used to programme those tables not generated by the general tableprogramme.

Without applying a general programme the number of hours is

$$O_2 = LT$$

If we want $O_1 \leq O_2$ we must have $t \leq pLT$

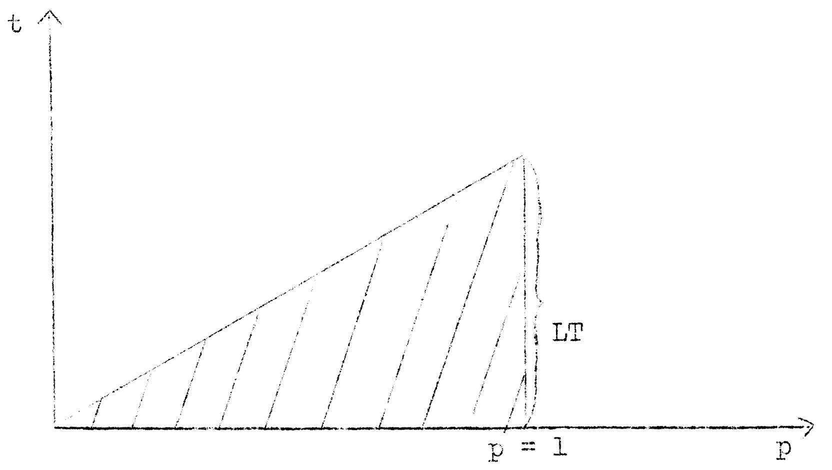


Fig. 1

The point (p,t) must thus belong to the shaded domain of fig. 1. From this is seen that if LTp is large, we can choose to develope a general programme even if we have to spend much time on the development.

General programmes for sorting and merging are examples on programmes for which LTp is large.

In this model we shall consider L and T to be exogenoues variables. Furthermore we assume that between p and t there exists a functional relationship expressed by

$$t = f(p) \qquad 0 \leq p \leq 1$$

It is difficult to specify the shape of f in general. It depends on the homogeneity of our tables, on the competence of the persons developing the general programme, and other factors left out of this model. A reasonable assumption is, however, that f is a function with non-negative first and second dorivatives.

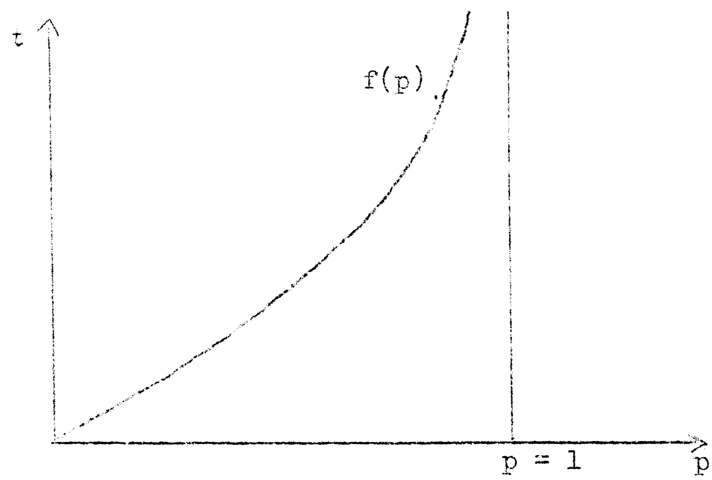


Fig. 2

2.1 minimalization of O_1

Inserting $t = f(p)$ in the expression for O_1 we get

$$O_1 = f(p) + (1 - p)\,LT$$

and

$$\frac{dO_1}{dp} = f'(p) - LT$$

in the minimum we have

$$f'(p) = LT$$

Graphically the minimum (p_o, t_o) is the intersection of two curves as showed in fig. 3.

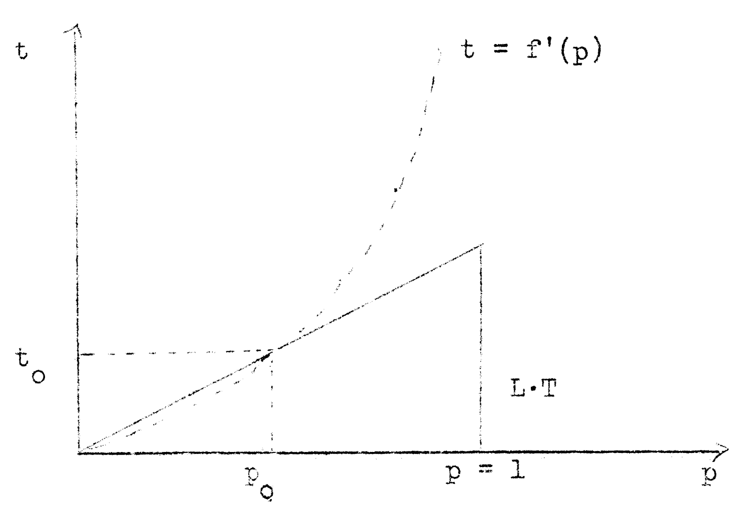


Fig. 3

From fig. 3 is seen that p_o and t_o will increase with LT.

2.2 An example

We assume that

$$f(p) = \frac{T \cdot p}{1-p}$$

which means that if we for instance use T hours to develope the general programme, then this programme generates 50% of our tables during the periode of L months. We get

$$f'(p) = \frac{T}{(1-p)^2}$$

which means that

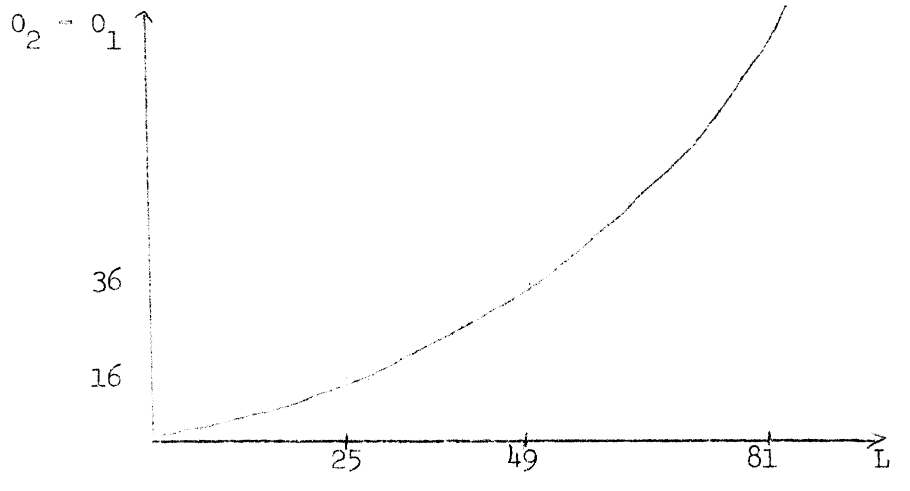
$$p_o = 1 - \frac{1}{\sqrt{L}}$$

Let L = 25 (about 2 years) and we will get the values

$$p_o = \frac{4}{5} \quad \text{and} \quad t_o = 4\ T$$

and the corresponding $0_1 = 9\ T$ and $0_2 = 25\ T$

In fig. 4 $0_2 - 0_1$ is shown as function of L



$$0_2 - 0_1 = LT + T - 2\ T\sqrt{L}$$

3. THE EFFECT OF THE GENERAL TABLEPROGRAMME ON THE PRODUCTIONTIME

By productiontime is here meant the time used on the computer to produce the tables.

Introduce

G: The number of hours used per. month on the computer to produce all tables when we do not apply the general tableprogramme.

a: The efficiency ratio defined as the ratio between the productiontimes per table with and without applying the general tableprogramme.

The productiontime with the general tableprogramme during the periode of L months then becomes

$$D_1 = pLGa + (1 - p)\ LG$$

Without the general tableprogramme the productiontime is

$$D_2 = LG$$

To get $D_1 \leq D_2$ we must have

$$p\ (a - 1) \leq 0$$

or

$$a \leq 1$$

This is probably a condition which is rarely met because generated programmes are likely to be less efficient than tailormade programmes.

The reason why this inequality is independent of L and G is that the time used on training users of the general programme is left out of this model. In other words, we assume that this amount of time is small as compared with D_1.

4. THE EFFECT OF THE GENERAL PROGRAMME ON THE TOTAL COST

Introduce

q: The price per. hour for programming.

r: The price per. hour for production.

The total cost when using the general programme is

$$C_1 = qt + (1 - p)TLq + rpLGa + (1 - p)LGr$$

When not using the general programme, the cost will be

$$C_2 = LTq + LGr$$

If we want $C_1 \leq C_2$ we must have

$$C_2 - C_1 = -qt + pLTq - rpLGa + pLGr > 0$$

If for instance $t = 0$ (which is the case if we take a general programme from a programme library) we must have

$$L \cdot p \quad (Tq + Gr - rGa) > 0$$

or

$$a < \frac{Tq + Gr}{rG} = 1 + \frac{Tq}{Gr}$$

The right side of this inequality is an increasing function of $\frac{Tq}{Gr}$. This means that if our tables requires relatively large programmingtime, we can choose a general tableprogramme from the library even when this programme has a bad efficiency ratio.

4.1 Optimalization

We have

$$U_1 = qt + (1 - p)\ TLq + rpLGa + (1 - p)\ LGr$$

and now t is a function of both p and a

$$t = g\ (a, p)$$

Inserting $g\ (a, p)$ in the expression for U_1 we can, subject to certain assumptions about the form of function g, minimize U_1 with respect to a and p, and determine the corresponding t.

This may be considered the long-range aim for a development during which the capasity of competent programmers can be adjusted to the need.

As to the shape of g, like f in section 2 it depends on the homogeneity of the tables we want to generate, on the competence of the persons developing the general programme and other factors left out of this model.

4.2 Minimalization for fixed t

In many cases a more restricted situation than the above described may be realistic in the short run. The capasity for developing general table-programmes, t may be determined by factors outside this model, and the problem will then be to choose a and p for this fixed value of t. We assume that for $t = t_o$ we have

$$a = h(p)$$

Here $h(p)$ is dependent of t_o, then

$$U_1 = qt_o + (1 - p)\, TLq + rpLGh(p) + (1 - p)\, LGr$$

$$\frac{dU_1}{dp} = - TLq + rLGh(p) + rLGph'(p) - LGr$$

In the minimum we have $\frac{dU_1}{dp} = 0$ or

$$h(p) + ph'(p) = \frac{Tq + Gr}{Gr} \text{ (independent of L)} = 1 + \frac{Tq}{Gr}$$

5. ACKNOWLEDGEMENT

I am grateful to Mr. S. Nordbotten and Mr. E. Aurbakken for help and encouragement during the preparation of this paper.