PROFESOR ZDZISŁAW H. HELLWIG. ON THE OCCASION OF HIS 85TH BIRTHDAY

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY
Nr 9 (15)

ISSN 1644-6739

Walenty Ostasiewicz
Uniwersytet Ekonomiczny we Wrocławiu

1. Brief biography

Prof. Dr. hab. h.c.h.c. Z.H. Hellwig was born on 26 May 1925 in a small town Dokszyce not for away from Wilno. Both of his parents were teachers. His father Henry Hellwig taught the German language, and his mother taught the mathematics. Prof. Z. Hellwig was educated at King Zygmunt August School in Wilno. The secondary school graduation certificate (matura, in Polish) he obtained however after the Second World War in Wrocław in 1947. He obtained his bachelor degree in Warsaw where he studied under the well-known Polish economist Professor Oskar Lange. In 1952 he received his Master Degree (magister, in Polish) at SGPiS (The School of Planning and Statistics) in Warsaw.

Being the student of the second year in the Higher Commercial School (Wyższa Szkoła Handlowa in Polish) he started to work at this School as a younger assistant.

On the basis of the work *Linear regression and its applications in economics* he got in 1958 his Ph.D. degree in economic sciences. The following year he came to England – at the invitation of Mr. Reddaway, Director of the Department of Applied Economics, University of Cambridge – and studied mathematical economics and statistics in London and Cambridge.

In 1967 he became the Professor in economics, and since 1972 he is the Ordinary Professor in economics.

In 1962 Prof. Z. Hellwig was nominated the Head of Department of Statistics and held this position till 1995, this is to the year when he retired.

Besides being the Head of Department, Professor Hellwig performed a number of duties. For three times he served as the vice-

Rector, he was also the Dean of the Faculty, editor-in-chief of the Research Papers of Wrocław School of Economics.

Professor Hellwig was the founding editor of the prestigious Polish journal *Operations Research and Decisions*, and is still active as a vice-chair of the Scientific Committee.

Professor Hellwig holds a number of awards, including Ministry awards, prestigious award by the Prime Minister of Polish Government, the Medal from Polish Statistical Society.

Among his many honors, he received an honorary doctorate from Cracow University of Economics in 1985, and from Prague University of Economics in 1994. He was elected as the honorary member of the Polish Academy of Science Committee for Statistics and Econometrics.

Professor Hellwig enjoys discussions, especially with his friends and young people friends. For hours long he is able to chat on any subject, including politics, economy, sex, mathematics, philosophy, and history. He was, and he still is, admired, respected and loved.

Professor Hellwig has authored or edited many books, and roughly 250 papers. The works on the variable selection and on the measure of socio-economic development are the most cited, and most influential papers in applied statistics.

2. Regression analysis

The problem of regression is probably the only scientific problem which Professor Hellwig studied during his whole scientific life. To this problem he devoted his doctoral thesis, it constituted the core of his habilitation dissertation, he published many papers on this subject. Professor Hellwig introduced a new definition of the regression concept. The new definition is given as follows.

Let a and b be any real numbers which satisfy the relation:

$$\int_{-\infty}^{+\infty}\int_{a}^{b} f(x, y)dx\,dy > 1-\alpha,$$

where $f(x, y)$ is a density function, and α is any small number from the interval [0, 1].

The regression is defined as a set R_k of functions $\psi(x, \alpha_1, \alpha_2, \ldots, \alpha_k)$, such that the following relation:

$$\underset{x\in[a,b]}{P}\left[\left|Y-\psi(x)\right|<\Delta\right]=\int_a^b\int_{\psi(x)-\Delta}^{\psi(x)+\Delta}\int f(xy)\,dxdy\geq 1-\alpha$$

is satisfied for some big number Δ. Simplified version of this relation is the following:

$$\underset{x\in[a,b]}{P}\left[\psi(x)-\Delta<Y<\psi(x)+\Delta\right]\geq 1-\alpha.$$

3. Hellwig's method of variable selection

In the paper on variable selection (see [Hellwig 1968]) the measure for evaluating the informational capacity of exogenous variables in econometric model was introduced.

Let $X_{(n)} = \{X_1, X_2, \ldots, X_n\}$ be a set of potentially possible variables which can be used for the prediction of another variable by means of linear predictor of the form:

$$Y = a_0 + a_1X_1 + \ldots + a_nX_n.$$

The problem is to determine the most informative m-element subset $X_{(m)} = \{X_{i_1}, X_{i_2}, \ldots, X_{i_m}\}$ of the given set $X_{(n)}$.

The crucial subproblem is the definition of the appropriate measure of the information conveyed by the chosen variables. The information measure conveyed by the m-element subsets $\{X_{k_1}, X_{k_2}, \ldots, X_{k_m}\}$ of the potential variables $\{X_1, X_2, \ldots, X_n\}$ has been defined by the following formula:

$$H(m,k,)=\sum_{i=1}^{m}\frac{r_{k_i}^2}{\sum_{j=1}^{m}|r_{k_ik_j}|},$$

where:

$k = (k_1, k_2, \ldots, k_m)$,

$r_{k_ik_j}$ – the correlation coefficient between X_{k_i} and X_{k_j},

r_{k_i} – the correlation coefficient between Y and X_{k_i}.

This measure turned out to be extremely useful in practical application.

4. Hellwig's measure of development

One of the biggest achievements of Prof. Hellwig is the famous, now the so-called Hellwig's measure of development. This measure has been defined for solving the following problem. Suppose there is given some finite set of N entities, such that countries or regions, which are characterized by n variables $X_1, X_2, \ldots, X_n$.

Let symbol x_{ij}, $i = 1, 2, \ldots, N$, $j = 1, 2, \ldots, n$ denote the value of the j-th variable for the i-th entity, or object. The problem is to determine the level of the economic development for each entity (object) under consideration. In order to solve this problem one has first of all to choose some reference point $x_0 = (x_{01}, x_{02}, \ldots, x_{0n})$, Hellwig called it "the pattern of economic development". Next, for each object one has to calculate its level of development according to the following formula:

$$d_i = 1 - \frac{c_i}{c_0}, \ i = 1,2,\ldots, N,$$

where N is the number of countries evaluated with respect to their status of development, and the quantities c_i and c_0 are defined as follows:

$$c_i = \left[\sum_{j=1}^{n} (x_{ij} - x_{0j})^2 \right]^{1/2},$$

$$c_0 = \bar{c} + 2\left[\frac{1}{N} \sum_{i=1}^{N} (c_i - \bar{c})^2 \right]^{1/2},$$

with $\bar{c} = \frac{1}{N} \sum_{i=1}^{N} c_i$.

Vector $x_i = (x_{i1}, x_{i2}, \ldots, x_{in})$ contains the n features charactering the evaluated countries and the reference vector $x_0 = (x_{01}, x_{02}, \ldots, x_{0n})$ has been called "the pattern of economic development".

ŚLĄSKI PRZEGLĄD STATYSTYCZNY
Nr 9 (15)

5. Measure of stochastic dependence

The other significant achievement of Prof. Hellwig is the measure of stochastic dependence. For the case of two-dimensional random vector (X, Y) this measure has been defined as follows:

$$d = (1 - \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \min[f(x,y), f_1(x)\cdot f_2(y)dxdy)^{1/2} .$$

For the discrete case this measure has been defined by the following formula:

$$d = \left(\frac{1 - \sum_{i=1}^{r}\sum_{j=1}^{s} \min(p_{ij}, p_i q_s)}{1 - (\min(r,s))^{-1}} \right)^{1/2} .$$

Both of them were further investigated in a number of papers.

6. Distance variable

In the book on stochastic approximation published in 1965 Z. Hellwig introduced a new statistical concept, and namely, the concept of distance variable [Hellwig 1965, 1969]. It has been defined as follows.

Let $X^0 = (X_1^0, X_2^0, \ldots, X_n^0), X^1 = (X_1^1, X_2^1, \ldots, X_n^1), \ldots,\ X^m = (X_1^m, X_2^m, \ldots, X_m^m\}$ be a simple random sample from the distribution given by cdf $F(x_1, \ldots, x_n)$ or by the density function $f(x_1, \ldots, x_n)$.
The distance random variable, denoted by symbol $C_{m,n}$, is defined as follows:

$$C_{m,n,} = \min(Y_1, Y_2, \ldots, Y_m),$$

where

$$Y_j = \left(\sum_{i=1}^{n} (X_i^0 - X_i^j)^2 \right)^{1/2} \quad j = 1, 2, \ldots, m.$$

ŚLĄSKI PRZEGLĄD STATYSTYCZNY Nr 9 (15)

Intuitively, variable $C_{m,n}$ means the shortest distance between a random vector X^0 and a set of random vectors $X_2^1, X^2, \ldots, X^m$.

The general expression for the cumulative distribution function found by B. Kopociński is following

$$F_{C_{m,n}}(c) = 1 - \int_{R^n} (1 - V(x_1, x_2, \ldots, x_n, c))^m \, dx, \ldots dx_n,$$

where:

$$V(x_1 \ldots x_n, c) = \int_A f(x_1 + u_1, x_2 + u_2, \ldots, x_n + u_n) \, du_1 \ldots du_n$$

with

$$A = \{(u_1 \ldots, u_n) \mid \sqrt{u_1^2 + u_2^2 + \ldots + u_n^2} < c\}.$$

The limit distribution, when $m \to \infty$, is following

$$F_{C_n}(c) = 1 - \int_{R^n} e^{-K_n(c)} f(x_1, x_2, \ldots, x_n) dx_1 \ldots dx_n,$$

where:

$$K_n(c) = \Pi^{n/2} c^n / \Gamma(n/2 + 1).$$

G. Trybuś reports the explicit exact formulae for these distributions for some populations in his monograph (*Zmienna losowa dystansowa. Teoria i zastosowania*, Prace Naukowe Akademii Ekonomicznej nr 173, AE, Wrocław 1981), where there are references to the original works of B. Kopociński and W. Dziubdziela.

7. Excerpt from UNESCO paper

The well known taxonomic method developed by Z. Hellwig has been first published by UNESCO. The fragment of the paper is given below.

5. VARIABLES CLUSTERING BY MEANS OF THE TAXONOMIC METHOD

To make the method more easily understandable we shall in the course of its presentation illustrate it with numerical examples.

(i) The algorithm starts with a correlation matrix R (n x n) where n stands for the number of variables (see Table 1[3]).

(ii) In the second step the following transformation is effected $d_{ij} = 1 - | r_{ij} |$ by means of which the matrix R is mapped into a new matrix D (see Table 2). This matrix is also of (n x n) type, symetrical and with elements d_{ij} where $i, j = 1, 2, \ldots$ and $d_{ij} \geqslant 0$)

(iii) Now the smallest element has to be found in each row of the matrix D. If we denote these elements by d_i then $d_i = \min_j d_{ij}$

(iv) After having found d_i in the i^{th} row, one finds the number of the column which forms with the given row a cross-cell containing the value d_i. In this way one gets n pairs of indices:

$$(1, j_1), (2, j_2), \ldots\ldots (n, j_n)$$

Let us now define a few terms. The terminology will be very simple. Each of our pairs of indices will be called a *link* whereas i and j_i in (i, j_i) are *nodes*. The first node is the *beginning* of the link which ends with the second node. The links may therefore be called oriented links. When we draw a graph all nodes will be represented by points and all links by arrows. The number of arrows pointing at a given node will be called the *power* of that node.

(v) The node with the highest power is next found. If there are more than one with the highest power the choice will fall on the one to which corresponds the smallest value of

$$\bar{d}_i = \frac{1}{n} \sum_{j=1}^{n} d_{ij}$$

(vi) Now all nodes which are beginnings of links pointing towards the node of the highest power are to be linked with this node. All nodes attached to the node of the highest power can be now looked upon as the ends of some other nodes which are in to be linked with all these ends. By continuing this process of joining together all the nodes related to the node of the highest power, one obtains a certain family of nodes which can be referred to as the "first concentration" of nodes, or as the first *cluster* of points.

(vii) Next, one finds all the clusters which correspond to the nodes with powers below the highest one, in a decreasing order of magnitude.

(viii) The node with the highest power in a cluster is termed the centre of the cluster. If in a cluster nodes of the highest order number more than one, we select the one to which corresponds the highest value of d_j. The centre of a cluster can be looked upon as the representative of variables belonging to this cluster. The set of representatives of all clusters will be referred to as a *set of core variables.*

(ix) The range of variation of the parameters d_i should now be divided into three parts:

---

1. Studies VI, VII, IX (Unesco mimeographed working documents).
2. A valuable example of selecting key indicators from a wide range of these is given in a research report of the United Nations Research Institute for Social Development: Report No. 70.10 Geneva 1970: Contents and measurement of socio-economic development: An empirical enquiry.
3. N.V. Sovani, A.M. Subramanian: An Index of Socio-Economic Development of Nations, UNRISD, 1969.

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY
Nr 9 (15)

$$d_i < \bar{d} - s_d, \ \bar{d} - s_d \leqslant d_i \leqslant \bar{d} + s_d; \ d_i > \bar{d} + s_d$$

where $\bar{d} - \frac{1}{n}\sum_{i=1}^{n} d_i$; $s_d = \left[\frac{1}{n} \sum_{i=1}^{n} (d_i - \bar{d})^2\right]^{\frac{1}{2}}$

All nodes which are beginnings of some links the length d_j of which belongs to the first interval can be removed from the graph and the corresponding variables may be called *redundant*. Nodes which are beginnings of some other links the length d_i of which falls in the third interval can also be eliminated and the respective variables may be referred to as *irrelevant*. All remaining variables would form the "compact" set. All the steps which lead to the obtention of particular clusters, the core variables as well as the "compact" set of variables are summarized and illustrated in the diagrams which follow.

The procedure which has been described is simple and liable to easy interpretation. It is equally applicable to variables which are continuous, discrete or categorial. If however one is put off by, or is suspicious of, its great simplicity, there are more sophisticated methods available, like for example, the principal components method, which will provide good solutions to the problem as it has been posed above, but at much greater cost.

6. CONCLUSION

The selection method which has been proposed, while being a variation on the intercorrelation theme, has the virtue of simplicity and provides a clear picture, through the taxonomic graph, of the relationships and distances between the variables. Furthermore a novel idea has found expression in the elaboration of the method: it refers to the problems of selection and weighting as close correlates requiring a unique solution.

The clustering technique which has been described here has identified clusters of variables and cluster centres. Each centre could be made representative of the whole group. Expanding the number of the more representative variables would mean a movement in concentric circles from the centre towards the periphery so as to include the number of variables which would be retained for any specific purpose. The two exercises which have been carried out have led to taxonomic graphs which are reproduced on the following pages. It may be that in each case the original number of variables has been somewhat too limited to derive many useful conclusions from the clustering. Most of the groupings appear to be consistent with what is already known or suspected of the relative movement of variables as development proceeds.

The distance matrix is following.

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY
Nr 9 (15)

TABLE II MATRIX D

i \ j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	$\bar{d}_i$
1	0	.56	.30	.70	.33	.31	.29	.43	.27	.28	.44	.25	.44	.29	.44	.47	.45	.21	.86	.30	.83	.74	.44	.43	.44
2		0	.36	.93	.70	.68	.56	.52	.67	.73	.73	.63	.77	.94	.87	.95	.64	.48	.95	.68	.86	.81	.68	.72	.71
3			0	.72	.52	.51	.33	.34	.35	.38	.46	.39	.47	.58	.53	.58	.53	.26	.93	.48	.64	.59	.57	.43	.49
4				0	.89	.93	.68	.85	.50	.71	.92	.67	.84	.68	.21	.34	.87	.65	.96	.71	.90	.88	.83	.95	.75
5					0	.10	.30	.39	.35	.25	.35	.27	.36	.35	.51	.51	.42	.21	.99	.27	.94	.83	.48	.29	.46
6					.10	0	.40	.44	.38	.31	.38	.32	.40	.41	.59	.57	.46	.26	.90	.27	.98	.94	.50	.32	.49
7							0	.25	.25	.22	.36	.27	.40	.33	.47	.52	.41	.14	.83	.28	.99	.84	.51	.34	.43
8							.25	0	.48	.44	.47	.43	.46	.52	.67	.66	.55	.30	.99	.49	.89	.79	.69	.42	.54
9									0	.12	.32	.17	.25	.46	.34	.33	.41	.21	.91	.38	.81	.72	.46	.28	.45
10									.12	0	.22	.18	.24	.40	.42	.46	.37	.20	.97	.33	.86	.73	.46	.21	.42
11											0	.29	.09	.55	.59	.58	.42	.33	.90	.42	1.0	.86	.62	.07	.49
12									.17			0	.28	.46	.37	.43	.23	.21	.89	.26	.85	.72	.43	.26	.40
13													0	.63	.52	.52	.42	.36	.87	.40	1.0	.85	.67	.06	.49
14	.29													0	.48	.48	.61	.31	.98	.38	.96	.93	.57	.58	.56
15															0	.12	.34	.40	.96	.31	.88	.79	.68	.58	.52
16																0	.08	.50	.99	.48	.79	.10	.82	.55	.51
17																.08	0	.44	.91	.41	.93	.93	.56	.43	.40
18							.14											0	.98	.24	.90	.78	.43	.29	.40
19																			0	.91	.60	.84	.90	.90	.91
20																		.24		0	.91	.85	.26	.36	.45
21																					0	.49	.79	1.0	.86
22																.10						0	.76	.78	.76
23																				.26			0	.59	.59
24													.06											0	.47

$\bar{d}_i = \frac{1}{n-1} \sum_{j=1}^{n} d_{ij}$

$S_d = \left| \frac{1}{n} \sum_{i=1}^{n} (d_i - \bar{d})^2 \right|^{1/2} = 14$

($d_i = \min d_{ij}$; $j = 1, \ldots n$)

$\bar{d} = \frac{1}{n} \sum_{i=1}^{n} d_i$

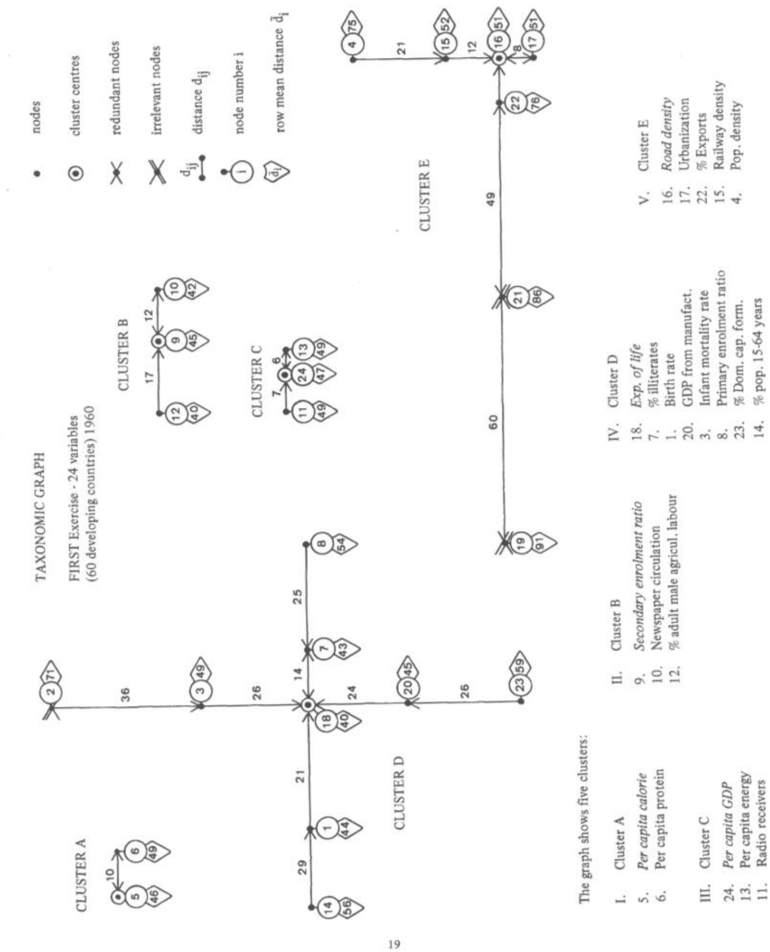
$= 19$

$d_i < 5$
(redundant)

$5 \leqslant d_i \leqslant 33$
(compact set of variables)

$d_i < 33$
(irrelevant)

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY
Nr 9 (15)



Selected works by Z. Hellwig

Hellwig Z., *Problem optymalnego wyboru predykant*, "Przegląd Statystyczny" 1968, No. 3–4.

Hellwig Z., *Aproksymacja. stochastyczna*, PWE, Warszawa 1965.

Hellwig Z., *On the measurement of stochastical dependence*, "Zastosowania Matematyki" 1969, Vol. 10, pp. 233–247.

Hellwig Z., *Efekt katalizy jego wykrywanie i usuwanie*, "Przegląd Statystyczny" 1977, No. 2, pp. 179–192.
Hellwig Z., *Przechodność skorelowania między zmiennymi losowymi i płynące stąd wnioski ekonometryczne*, "Przegląd Statystyczny" 1976, No. 1, pp. 3–20.
Hellwig Z., *Regresja liniowa i jej zastosowanie w ekonomii*, PWE, Warszawa 1960.
Hellwig Z., *Linear Regression and Its Application to Economics*, Pergamon Press, Oxford 1963.
Hellwig Z., *Przyczynek do teorii regresji*, Zeszyty Naukowe WSE, Wrocław 1961.