Rank-Correlation among the Four Class-Records in a College

SINGU, Tadao

(Laboratory of Mathematics and Statistics)

大学生の成績の順位相関について

新　宮　忠　雄

（数学および統計学研究室）

§ 1. Introduction

The effect of the entrance examination to those who apply a college is often discussed in our country, but we have few scientific data for this purpose. Our president, Dr. Takezaki, Yosinori suggested me to gather such data from the graduates of the Simane Agricultural College and to analyze them. In March of 1958 our college sent off the fourth-time graduates, so even only in our college still we can get twelve kinds of data for this purpose, for there are three classes each year.

To calculate the correlation coefficient between the grades at the entrance examination and those at the graduation one may think that it would be better to use the average scores of each student than to use the ranks in his class. Yes, that may be true, but its calculation is very tedious and yet the results are almost the same. See one example done for the data from (F2) at p. 144. Therefore I decided to use their ranks in the class instead of the average scores. In selecting the sample I did as follows:

Those who joined the class later or earlier were omitted from the ranking. It means that if a student was admitted into his class at the beginning of the second or third year from other college he was eliminated from the ranking. And also if a student could not graduate after four years by his absence from school or by the failure in the examination he was omitted from the member of his class. The reason for such elimination is simple. Students in one class should had the same or at least almost the same examinations at the entrance and in the college education.

At the end of the introduction I wish to express my great appreciation to Miss Takeda, Yukiko for her gathering data from the students' records.

§ 2. Data

Table 1. Notation for Each Class.

Entered in	Graduated in	Name of the Class		Notation
		Course	Time	
1951	1955	A	1st	A1
〃	〃	F	1st	F1
〃	〃	E	1st	E1
1952	1956	A	2nd	A2
〃	〃	F	2nd	F2
〃	〃	E	2nd	E2
1953	1957	A	3rd	A3
〃	〃	F	3rd	F3
〃	〃	E	3rd	E3
1954	1958	A	4th	A4
〃	〃	F	4th	F4
〃	〃	E	4th	E4

A : Agriculture

F : Forestry

E : Agriculture and forestry economics

Other notations are as follows:

X: Rank of the total score at the entrance examinations,
Y: Rank of the average score in the whole course of college education,
Z: Rank of the average score in the course of general education,
V: Rank of the average score in the course of professional education.

Table 2. Data from (A1)

X	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Y	7	6	1	4	5	2	13	3	9	17	19	21	15	22	16	12	10	11	8	14	18	20
Z	4	3	1	7	6	5	13	8	9	15	22	12	16	18	17	2	10	19	11	14	20	21
V	9	6	3	4	5	2	19	1	10	18	17	22	13	21	15	14	12	8	7	11	16	20

Table 3. Data from (F1)

X	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Y	9	1	12	4	16	13	5	17	15	20	3	11	10	19	18	2	14	7	8	6
Z	6	1	12	3	9	8	4	13	15	19	14	16	17	20	11	2	18	7	5	10
V	13	2	12	4	16	14	6	17	15	20	1	10	9	18	19	3	8	7	11	5

Table 4. Data from (E1)

X	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Y	9	1	12	15	13	5	4	6	16	17	20	11	26	8	25	3
Z	7	4	16	10	12	2	1	3	11	21	15	13	19	6	28	8
V	12	1	11	16	13	6	5	9	17	14	22	10	26	8	23	3

X	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
Y	28	29	21	22	19	23	14	31	24	10	2	7	30	27	18
Z	22	27	14	25	24	17	5	20	26	29	9	23	31	30	18
V	29	30	24	20	15	25	19	31	21	7	2	4	28	27	18

Table 5. Data from (A2)

X	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Y	3	1	2	11	4	7	10	8	18	6	9	5	19	14	12	16	15	17	13	20
Z	2	1	4	8	6	9	10	3	12	7	11	5	20	16	13	19	14	17	15	18
V	3	1	4	12	2	5	8	10	18	6	9	7	19	14	11	15	17	16	13	20

Table 6. Data from (F2)

X	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Y	1	12	7	5	4	13	3	2	22	9	21	16	17	8	18	10	20	15	6	23	19	14	11
Z	1	13	9	7	4	17	3	2	19	16	22	11	18	8	14	10	20	15	5	23	21	12	6
V	1	12	7	5	3	9	2	4	23	6	20	19	16	10	18	8	21	14	11	22	17	15	13

Table 7. Data from (E2)

X	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
Y	2	5	8	4	3	1	6	11	21	13	16	9	10	14.5	14.5	18	7	20	19	17	12	24	22	25	26	27	28	23
Z	2	5	13	4	3	1	10	6	20	8	11	9	12	15	23	21	7	18	16	17	14	22	24	25	26	28	27	19
V	2	6	7	5	3	1	4	17	25	14	12	9	8	16	10	15	13	19	20	18	11	26	24	21	22	27	28	23

Table 8. Data from (A3)

X	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20.5	20.5	22.5	22.5	24	25
Y	3	7	6	8	2	10	11	12	1	14	5	20	4	16	17	9	15	23	22	18	21	25	19	24	13
Z	2	8	16	7	3	6	11	9	1	12	4	19	5	18	15	10	14	23	21	13	22	25	20	24	17
V	5	7	4	8	2	12	9	14	1	16	11	20	3	13	17	6	15	22	23	21	19	25	18	24	10

Table 9. Data from (F3)

X	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Y	2	1	11	10	3	7	9	6	17	18	19	8	16	12	14	15	13	4	5
Z	10	2	17	5	1	4	8	11	18.5	16	18.5	7	15	9	14	12	13	6	3
V	2	1	3	10	7	11	9	5	15	18	19	8	17	14	12	16	13	4	6

Table 10. Data from (E3)

X	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15.5	15.5
Y	4	6	2	8	16	1	12	14	3	13	5	11	9	10	15	7
Z	4	5	1	7	9	2	6	12	3	13	8	14	15	10	16	11
V	5	7	2	8	16	1	15	14	3	11	4	10	9	12	13	6

Table 11. Data from (A4)

X	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
Y	4	7	21	6	15	8	13	18	12	17	9	5	11	3	27	20	2	22	1	16	26	14	24	25	23	10	19
Z	4	3	21	9	5	8	16	11	12	13	7	2	20	6	26	22	10	18	1	17	27	19	24	25	14	15	23
V	3	12.5	21	4	22	14	8	20	12.5	18	16	5	6	2	27	19	1	25	7	15	17	9.5	26	24	23	9.5	11

Table 12. Data from (F4)

X	1	2	3	4	5	6	7	8	9	10	11	12.5	12.5	14	15	16	17	18
Y	15	5	9	12	17	2	6	4	1	14	11	7	3	13	10	18	8	16
Z	10	6	9	15	16	3	5	1	4	14	11	12	8	2	17	18	7	13
V	15	5	9	11	18	2	6	4	1	14	12	7	3	13	10	16	8	17

Table 13. Data from (E4)

X	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Y	6	12	11	5	1	18	17	10	19	8	9	22	4	16	15	3	7	14	21	13	2	20
Z	6	13	8	4	5	16.5	22	7	20	9	11	19	15	18	12	2	3	10	21	16.5	1	14
V	5	11	12	6	1	19	7	10	17	8	9	22	3	16	18	4	13	15	21	14	2	20

§ 3. Calculation of Correlation Coefficients

Test statistic for the criterion by the rank-correlation coefficient is

$$r_s = 1 - \frac{6 \sum_{i=1}^{N} d_i^2}{N(N-1)}$$

where N is the number of pairs of observations of ranks (X_i, Y_i); $i = 1, 2, \ldots, N$; $d_i = X_i - Y_i$; (2), (3). When $X_i = X_j$, I ranked both of them as $\frac{1}{2}(X_i + X_j)$.

Table 14. $\sum_{i=1}^{N} d_i^2$

Class	N	$\Sigma(X-Y)^2$	$\Sigma(X-Z)^2$	$\Sigma(X-V)^2$	$\Sigma(Y-Z)^2$	$\Sigma(Y-V)^2$	$\Sigma(Z-V)^2$
A 1	22	580	584	922	352	91	610
F 1	20	1288	986	1440	388	74	656
E 1	31	2886	1908	3222	1280	132	1992
A 2	20	302	226	304	112	32	190
F 2	23	1005	1308	876	160	88	404
E 2	28	502.5	632	674	245.5	191.5	608
A 3	25	636	691	802	190	104	468
F 3	19	746	977.5	666	199.5	112	487.5
E 3	16	435.5	178.5	510.5	164	26	252
A 4	27	1994	1544	2620	536	444	1544
F 4	18	797.5	695.5	785.5	286	8	308
E 4	22	1450	1603.5	1230	294.5	158	660.5

Table 15. Rank-correlation Coefficients Obtained

Class	$r_{\text{cr. v.}}$		r_{XY}	r_{XZ}	r_{XV}	r_{YZ}	r_{YV}	r_{ZV}	N
A 1	0.508	0.359	0.672 (3) **	0.670 (4) **	0.479 (6) *	0.801 (2) **	0.949 (1) **	0.656 (5) **	22
F 1	0.534	0.377	0.032 (5) ◎	0.259 (4) ◎	−0.083 (6) −◎	0.708 (2) **	0.944 (1) **	0.507 (3) *	20
E 1	0.425	0.301	0.418 (5) *	0.615 (3) **	0.350 (6) *	0.742 (2) **	0.973 (1) **	0.598 (4) **	31
A 2	0.534	0.377	0.773 (5) **	0.830 (4) **	0.771 (6) **	0.916 (2) **	0.976 (1) **	0.857 (3) **	20
F 2	0.496	0.351	0.503 (5) **	0.354 (6) *	0.567 (4) **	0.921 (2) **	0.957 (1) **	0.800 (3) **	23
E 2	0.448	0.317	0.862 (3) **	0.827 (5) **	0.816 (6) **	0.933 (2) **	0.948 (1) **	0.834 (4) **	28
A 3	0.475	0.336	0.755 (4) **	0.734 (5) **	0.692 (6) **	0.927 (2) **	0.960 (1) **	0.820 (3) **	25
F 3	0.549	0.388	0.346 (5) ◎	0.143 (6) ◎	0.416 (4) *	0.825 (2) **	0.902 (1) **	0.572 (3) **	19
E 3	0.601	0.425	0.360 (5) ◎	0.738 (3) **	0.249 (6) ◎	0.759 (2) **	0.962 (1) **	0.629 (4) **	16
A 4	0.456	0.323	0.391 (5) *	0.529 (4) *	0.200 (6) ◎	0.836 (2) **	0.864 (1) **	0.529 (3) **	27
F 4	0.564	0.399	0.177 (6) ◎	0.282 (4) ◎	0.189 (5) ◎	0.705 (2) **	0.992 (1) **	0.672 (3) **	18
E 4	0.508	0.359	0.181 (5) ◎	0.095 (6) ◎	0.306 (4) ◎	0.834 (2) **	0.911 (1) **	0.627 (3) **	22

Expalanation for the table 15 :

	In the case of one-tail test	In the case of two-tail test
$r_{\text{cr. v.}}$ { of the left column	$\alpha = 1\%$	$\alpha = 2\%$
$r_{\text{cr. v.}}$ { of the right column	$\alpha = 5\%$	$\alpha = 10\%$
** : strongly correlated (positively)	$\alpha > 5\%$	$\alpha > 10\%$
* : weakly correlated (positively)	$1\% < \alpha < 5\%$	$2\% < \alpha < 10\%$
◎ : independent (positively)	$\alpha < 1\%$	$\alpha < 2\%$
−◎ : independent (negatively)	〃	〃

(p) means the value of "r" be the p-th among the six values of each class.

α denotes the probability of the error of the first kind.

Calculation of the Correlation Coefficient between Z and V by the Data from (F2)

Classes for V		65−70	70−75	75−80	80−85	85−90	f_z	$z\sum v$	$f_z z$	$f_z z^2$
Z	z \\ v	−2	−1	0	1	2				
65−70	−1	1	3	2			6	5	−6	6
70−75	0		2	9	1		12	0	0	0
75−80	1			1	1	2	4	5	4	4
80−85	2					1	1	4	2	4
f_v		1	5	12	2	3	N=23	14	0	14
$v\sum z$		2	3	0	1	8	14			
$f_v\ v$		−2	−5	0	2	6	1			
$f_v\ v^2$		4	5	0	2	12	23			

N.B. v=code of V
z=code of Z
f=frequency

If we assume the population to be binormal, the regressions to be linear and the variances to be homogeneous, the direction coefficients (slopes) of regression lines are as follows :

$$b=\frac{\Sigma(ZV)-\dfrac{\Sigma Z\ \Sigma V}{N}}{\Sigma V^2\ -\dfrac{(\Sigma V)^2}{N}}=\frac{14\ -\dfrac{1\times0}{23}}{23\ -\dfrac{1\times1}{23}}=0.621,$$

$$b'=\frac{\Sigma(ZV)-\dfrac{\Sigma Z\ \Sigma V}{N}}{\Sigma Z^2\ -\dfrac{(\Sigma Z)^2}{N}}=\frac{14\ -\dfrac{1\times0}{23}}{14\ -\dfrac{0\times0}{23}}=1.000.$$

So that we get as the correlation coefficient

$$r=\sqrt{bb'}=0.788.$$

By the David's chart[1], we have $0.55\leqq\rho\leqq0.90$, and $0.43\leqq\rho\leqq0.93$ as the confidence intervals of ρ, where their confidence levels are 95% and 99% respectively. Since both intervals do not contain $\rho=0$, we can conclude that Z and V are not independent. Here we find that the rank-correlation coefficient r_{ZV} (=0.800) is almost equal to the correlation coefficient r (=0.788).

§4. Discussion on the Rank-correlation Coefficients

Analization of the Table 16 :

The grand total of p's for each rs shows that r_{XY}, r_{XV} and r_{XV} seem very much larger than other three, or extremely larger than r_{YZ} and r_{YV}. To discuss such differences statistically we can use the criteria for testing for extreme mean[2],[3] or the method of analysis of variance. If we use the former its results are as follows :

For $k=6$ our test statistic is $r_{10}=\dfrac{\overline{X}_2-\overline{X}_1}{\overline{X}_k-\overline{X}_1}$, where

Table 16. Totals and Averages of p's

		r_{XY}	r_{XZ}	r_{XV}	r_{YZ}	r_{YV}	r_{ZV}	size of sample
grand total		56	54	65	24	12	41	12
grand mean		4.67	4.50	5.42	2.00	1.00	3.42	
total of	1's	13	11	18	6	3	12	3
〃	2's	13	15	16	6	3	10	3
〃	3's	14	14	16	6	3	10	3
〃	4's	16	14	15	6	3	9	3
〃	1,3 & 4's	43	39	49	18	9	31	9
〃	A's	17	17	24	8	4	14	4
〃	F's	21	20	19	8	4	12	4
〃	E's	18	17	22	8	4	15	4
average of	1's	4.33	3.67	6.00	2.00	1.00	4.00	
〃	2's	4.33	5.00	5.33	2.00	1.00	3.33	
〃	3's	4.67	4.67	5.33	2.00	1.00	3.33	
〃	4's	5.33	4.67	5.00	2.00	1.00	3.00	
〃	1,3 & 4's	4.78	4.33	5.44	2.00	1.00	3.44	
〃	A's	4.25	4.25	6.00	2.00	1.00	3.50	
〃	F's	5.25	5.00	4.75	2.00	1.00	3.00	
〃	E's	4.50	4.25	5.50	2.00	1.00	3.75	

$\overline{X}_1=5.42$, $\overline{X}_2=4.67$, $\overline{X}_k=1.00$ for the hypothesis H: $\bar{p}(r_{XV})$ is significantly larger than other $\bar{p}'$ s. We get $r_{10}=0.170$, and this value is smaller than its critical value 0.560 for 5% level of significance. Therefore we accept H. Also for the hypothesis H': $\bar{p}(r_{YV})$ is significantly smaller than other $\bar{p}'$ s, we have $\overline{X}_1'=1.00$, $\overline{X}_2'=2.00$, $\overline{X}_k'=5.42$,

$r_{10}'=0.227<0.560$.

Therefore we accept H', too.

As for other totals of p' s the results are the same to this, so I will omit the discussion for this table anymore. Instead of this I will use the method of analysis of variance applied to the table 15.

Let us separate r_s' s into two groups, namely ($h=1$)-group (r_{XY}, r_{XZ}, and r_{XV}) and ($h=2$)-group (r_{YZ}, r_{YV} and r_{VZ}). The year of graduate-time 1, 2, 3 and 4 may correspond to $i=1$, 2, 3 and 4 respectively. The courses A, F and E be denoted by $j=1$, 2 and 3 respectively. Now, there are three variables of classification with three observations per cell. According to the formula[4] the details of the analysis of variance are as follows:

Assumptions: Let the population be normal, i.e. $N(\mu_{hij}, \sigma_{hij}^2)$.

$\sigma_{hij}^2=\text{constant}=\sigma^2$ for any h, i and j.

All effects for the μ_{hij} are additive, i. e.

$\mu_{hij}=\mu+r_h+s_i+t_j+I_{hi}+I_{hj}+I_{ij}+I_{hij}$,

where μ is the common value to each cell, r_h is group-effect, s_i is year-effect, t_j is course-effect, and I_{hi}, I_{hj}, I_{ij}, I_{hij} are all interactions between these effects.

Assume the sample to be random.

Hypotheses: H_1: there is no group (R)-effect, i.e. $r_h = 0$ $(h=1, 2)$.

H_2: no year (S)-effect, i.e, $s_i = 0$ $(i=1, 2, 3, 4)$.

H_3: no course (T)-effect, i.e. $t_j = 0$ $(j=1, 2, 3)$.

H_4: no RS-interaction, i.e. $I_{hi} = 0$ $(h=1, 2; i=1, 2, 3, 4)$.

H_5: no RT-interaction, i.e. $I_{hj} = 0$ $(h=1, 2; j=1, 2, 3)$.

H_6: no ST-interaction, i.e. $I_{ij} = 0$ $(i=1, 2, 3, 4; j=1, 2, 3)$.

H_7: no RST-interaction, i.e. $I_{hij} = 0$ $(h=1, 2; i=1, 2, 3, 4; j=1, 2, 3)$.

Notations: X_{hijl} = observed value,

$h=1, 2(=r); i=1, 2, 3, 4(=s); j=1, 2, 3(=t); l=1, 2, 3(=m)$.

$$T_{h\cdots} = \sum_{i=1}^{s}\sum_{j=1}^{t}\sum_{l=1}^{m} X_{hijl}, \quad T_{.i..} = \sum_{h=1}^{r}\sum_{j}\sum_{l} X_{hijl}, \quad T_{..j.} = \sum_{h}\sum_{i}\sum_{l} X_{hijl},$$

$$T_{hi..} = \sum_{j}\sum_{l} X_{hijl}, \quad T_{h.j.} = \sum_{i}\sum_{l} X_{hijl}, \quad T_{.ij.} = \sum_{h}\sum_{l} X_{hijl},$$

$$T_{hij.} = \sum_{l} X_{hijl}, \quad T_{....} = \sum_{h}\sum_{i}\sum_{j}\sum_{l} X_{hijl}.$$

$$S_1 = \frac{1}{mst}\sum_{h} T_{h...}^2 - \frac{T_{....}^2}{N},$$

$$S_2 = \frac{1}{mrt}\sum_{i} T_{.i..}^2 - \frac{T_{....}^2}{N},$$

$$S_3 = \frac{1}{mrs}\sum_{j} T_{..j.}^2 - \frac{T_{....}^2}{N},$$

$$S_4 = \frac{1}{mt}\sum_{h}\sum_{i} T_{hi..}^2 - \frac{1}{mst}\sum_{h} T_{h...}^2 - \frac{1}{mrt}\sum_{i} T_{.i..}^2 + \frac{T_{....}^2}{N},$$

$$S_5 = \frac{1}{ms}\sum_{h}\sum_{j} T_{h.j.}^2 - \frac{1}{mst}\sum_{h} T_{h...}^2 - \frac{1}{mrs}\sum_{j} T_{..j.}^2 + \frac{T_{....}^2}{N},$$

$$S_6 = \frac{1}{mr}\sum_{i}\sum_{j} T_{.ij.}^2 - \frac{1}{mrt}\sum_{i} T_{.i..}^2 - \frac{1}{mrs}\sum_{j} T_{..j.}^2 + \frac{T_{....}^2}{N},$$

$$S_7 = \frac{1}{m}\sum_{h}\sum_{i}\sum_{j} T_{hij.}^2 - \frac{1}{mt}\sum_{h}\sum_{i} T_{hi..}^2 - \frac{1}{ms}\sum_{h}\sum_{j} T_{h.j.}^2 - \frac{1}{mr}\sum_{i}\sum_{j} T_{.ij.}^2$$
$$+ \frac{1}{mst}\sum_{h} T_{h...}^2 + \frac{1}{mrt}\sum_{i} T_{.i..}^2 + \frac{1}{mrs}\sum_{j} T_{..j.}^2 - \frac{T_{....}^2}{N},$$

$$S_8 = S_9 - (S_1 + S_2 + S_3 + S_4 + S_5 + S_6 + S_7),$$

$$S_9 = \sum_{h}\sum_{i}\sum_{j}\sum_{l} X_{hijl}^2 - \frac{T_{....}^2}{N}.$$

Table 17. Analysis of Variance

Source	Sum of squares	d. f.	Mean square	F-ratio
Groups	$S_1=2.292654$	1	2.292654	$115.0=F_1$
Years	$S_2=0.831497$	3	0.277166	$13.90=F_2$
Courses	$S_3=0.504696$	2	0·252348	$12.65=F_3$
RS-interaction	$S_4=0.223091$	3	0.074364	$3.73=F_4$
RT-interaction	$S_5=0.296717$	2	0.148358	$7.44=F_5$
ST-interaction	$S_6=0.194037$	6	0.032339	$1.62=F_6$
RST-interaction	$S_7=0.086289$	6	0.014365	$0.72=F_7$
Residual	$S_8=0.957159$	48	0.019941	
Total	$S_9=5.386140$	71		

$F_1>F_{0.995}(1,48)=8.7$ H_1 is rejected at the 0.5% level of significance.
$F_2>F_{0.995}(3,48)=4.9$ H_2 is rejected at the 0.5% level of significance.
$F_3>F_{0.995}(2,48)=5.9$ H_3 is rejected at the 0.5% level of significance.
$F_4>F_{0.975}(3,48)=3.4$ H_4 is rejected at the 2.5% level of significance.
$F_4<F_{0.99}(3,48)=4.2$ H_4 is accepted at the 1% leval of significance.
$F_5>F_{0.995}(2,48)=5.9$ H_5 is rejected at the 0.5% level of significance.
$F_6<F_{0.90}(6,48)=1.89$ H_6 is accepted at the 10% level of significance.
$F_7<F_{0.50}(6,48)=0.90$ H_7 is accepted at the 50% level of significance.

Conclusions: The group-effect is most significant, and the year-effect, the course-effect are second significant. Next comes the group-course-interaction. These are extremely significant, but the group-year-interaction is weakly significant, and other two interactions, namely, the year-course-interaction and the triple interaction, are not significant.

§ 5. Conclusion

We found that almost the half of the correlations between the remarks in the entrance examination and those at the graduation are independent, though almost all of the correlations between the three kinds of remarks at the graduation are closely correlated. It means that X is independent from Y, Z and V; and also V is not much correlated with Z though much correlated with Y. The reason of this is that in the professional education quite a few subjects are contained compared with those in the general education.

X is closely correlated to other grades without exception in the case of the second-time graduates. For the reason of this I can point out the fact that only in this year we examined two subjects from each of mathematics, natural sciences and social studies, though in the other years we examined only one subject from each of these. Of course, as for the weight compared with other subjects, namely, Japanese and English, were kept constant every year.

In the course of agriculture the rank-correlation coefficient between X and Y, Z or

V is always high, although in the class of forestry the coefficient is rather of low value. As for the course of agriculture and forestry economics it is almost the same to that of the course of forestry.

Judging from these facts, especially from the first result, I conclude as follows:

If we wish to choose better students by the entrance examination we have to examine at least two subjects from each of mathematics, natural sciences and social studies. If we cannot do so, the effect of the entrance examination is very little. Instead of such poor examination I would suggest to abolish them and to accept all students who applied, and after one or two semesters we can decide the students to be left in school according to the grades in these semesters. By this method we can select the best group of students for our college. Of course, there are many difficulties to accept all applicants without omission. For example, the shortage of classrooms and experimental apparatus is one of them. One solution for this is to postpone the subjects which contain experiments to later semesters and to lecture as many as one teacher can. By this reason this method would be rejected by almost all teachers if the number of them is not sufficiently large as in many state colleges in the U. S. A.

References

(1) David, F. N.: Tables of the Ordinates and Probability Integral of the Distribution of the Correlation Coefficient in Small Samples, The Biometric office, London.

(2) Dixon, W. J.: "Analysis of extreme values," *Annals of Mathematical Statistics*, 21, 488-506, 1950.

(3) Dixon, W. J. and Massey, F. J. Jr.: Introduction to Statististical Analysis; McGraw-Hill; New York, Toronto & London; 145-147, 261; 1951.

(4) Mood, A.McF.: Introduction to the Theory of Statistics; McGraw-Hill; New York, Toronto & London; 337-339; 1950.

(5) Olds, E. G.: "The 5% significance levels for sums of squares of rank differences and a correlation," *Annals of Mathematical Statistics*, 20, 117-118, 1949.

摘　　要

島根農科大学の第1期から第4期までの卒業生について各科（農学，林学，農林経済学）各期毎に，入学試験，一般教育，専門教育，卒業成績の間の相関をしらべた．普通の相関係数も1例について調べてみたが，標本の大きさが相当大きいところから，順位相関係数とほとんど同じ結果を示すので後者によった．

その結果を概観すると，入試に数学，理科，社会の各各から1科目ずつ選択させた年は，在学中の成績との相関は弱く，2科目ずつ選択した年は強い．またいずれの年でも，在学中の3種の成績間には相当強い相関関係があるのに，入試成績とはどれもほとんど相関していない．特に専門教育とは最もへだたっている．これは入試科目が一般教育の科目と同種のものであるのに反し，専門教育は職業教育である関係と思われる．また卒業成績は一般教育とよりも専門教育とより強く相関している．これは専門教育の方に学科数が圧倒的に多いためである．

結論として最も望ましい入学者決定方法としては，1〜2学期間全員を入学させ，その間の在学中の成績によるものが考えられる．しかしこれは多くの困難を伴なうから，やむをえず入試を行なうときは，少くとも2科目選択制をとらねばならない．

しかし何といっても資料が十分とはいゝかねるので，今後の調査研究にまつところが多い．