STATISTICAL QUALITY CONTROL IN JAPANESE INDUSTRY AND EDUCATION PROGRAMS FOR ENGINEERS

Tadakazu Okuno
Dean, Faculty of Engineering
Science University of Tokyo
Japan

1. Introduction

Development of quality control

Modern or Statistical Quality Control (SQC) took the first form in 1930s in the U.S. when W. A. Shewhart of Bell Telephone Laboratory applied the Control Chart invented by him to American industries. The modern statistical theories originated in the U.K. and the application of them to SQC was attempted as early as 1935. In Japan, the study of the modern statistics was engaged before the War by a limited number of researchers. The U.S. Armed Forces stationed in Japan after the War were concerned with the low and erratic quality characteristics of Japanese communication instruments and facilities.

From around 1950, SQC utilizing the Control Chart and the Sampling Inspection came to be used in many Japanese plants after the line of Dr. W. E. Deming. QC was first engaged only by engineers and factory workers. The original concept of QC was such that the conformity of quality could be assured through inspection, which is referred to as the Inspection-oriented QC.

If goods of non-conformity result continually from the production process, a rigid inspection process would become uneconomic. It is to be preferred to reduce the number of goods of non-conformity by means of eliminating negative factors in the production process. This stage of QC is called the Process-control-oriented QC.

Problems concerning reliability, safety and economy of products as well as of failing design and unfit materials cannot be solved only by the QC efforts of the production department. These problems are to be dealt with by a comprehensive QC containing the phases of development, plan and design of new products. This stage is called the Design-oriented QC.

Features of Japanese quality control

The present Japanese QC are represented in the following six items. Recently, some enterprises in other countries follow the similar lines of QC.

1) Company-Wide Quality Control (CWQC), which indicates the QC with all the departments and employees of a corporation participated;

2) QC with emphasis on education and training for all the employees from top managers to workers;

3) Quality Control with QC circle activities;

4) Quality Control with QC audits, by the president of the concerned corporation and by the Deming Prize Committee, whose members are composed of academic and research specialists from non-profit organizations;

5) QC with effective utilization of statistical methods;

6) Promotion of QC activities by the way of nation-wide movements.

In this paper we shall deal with Items 2) and 5) and illustrate several cases in automotive industries.

2. Quality Control Education Scheme

Very few educational programmes have ever been devoted to rearing QC experts in Japan. Presently, consultation specialists in the QC have had the educational backgrounds of chemical, mechanical, electric, electronic or industrial engineering. Only a small number of them have majored in mathematics or mathematical statistics.

Table 1. Main Education Courses for 1986 of JUSE

Courses	Term(t)		Enrol./t	Freq./yr
	days	hours		
[QC in general]				
1. Comprehensive Basic Course	30	200	150	7
2. Elementary Basic Course	8	50	150	6
3. QC Top Management Course	4	29	65	5
4. QC Executives Course	4	32	98	6
5. QC Middle Management Course	12	78	135	9
6. Basic Course for Foremen	6	41	110	15
7. TQC Instructor Course	6	48	135	4
8. QC-Circle Instructor Course	6	60	110	20
9. QC-Circle Leader Course	3	26	110	39
[Specific Topics]				
1. Design of Experiment (DE)	30	189	100	2
2. Elementary Course in DE	8	54	140	7
3. Elementary Multivariate Analysis	4	24	100	3
4. Advanced Multivariate Analysis	3	18	90	1
5. Reliability	15	103	100	3
6. Elementary Reliability	4	28	100	7
7. Management Course in Reliability	4	24	100	3
8. FMEA-FTA	2	13	130	11
9. Design Review	3	18	100	8
10. Operations Research	30	190	50	1
11. Marketing Research	17	110	90	1
12. Sensory Inspection	11	67	60	1

A number of QC education courses are held every year for staff members of all ranks of enterprises by institutions like JUSE, JSA etc. As an example, the education courses for 1986 of JUSE are shown in Table 1.

Table 2. Curriculum for Design of Experiment Course

Subject	No. of Units*	Subject	No. of Units
0. Preliminary Course (2)		5. Techniques for Planning and Designing Products (11)	
Statistical Methods	2	DE for designing	3
1. Introduction (4)		Quality Deployment System	3
Basic Concept	1	FMEA/FTA	3
Factor and Level	1	Analysis of Process Capability	2
DE in Production Process	1	6. Specific Topics (9)	
Use of DE	1	Signal-to-Noise Ratio	2
2. Design and Analysis of Experiments (22)		Multivariate ANOVA	1
Multi-way Layouts	5	Use of Micro-computer	1
Split-plot Designs	2	Analysis of Sensory Inspection Data	1
Use of Orthogonal Arrays	9	Exploratory Data Analysis	1
Incomplete Block Design	2	Practical Examples	2
Composite Designs	2	Special Lecture	1
Analysis of Categorical Data	2	7. Exercises and Others (11)	
3. Use of Supplementary Variables (3)		DE Game (Group Competition)	2.5
Regression Analysis	3	Exercises	5
4. Theory (3)		Group Discussions	3
Linear Statistical Inference	3	Concluding Remarks	0.5

* A unit denotes 3-hour lecture and/or practice.

Table 3. Curriculum for Elementary and Advanced Multivariate Analysis (MA) Courses

Elementary MA course: Subject	No. of Units	Advanced MA course: Subject	No. of Units
What is MA	1	Multiple regression and PCA	1
Simple regression	1	Discriminant Analysis	0.5
Multiple regression (MR) ($p^* = 2$)	1	Cluster Analysis	1
Principal Component Analysis (PCA)	1	Multi-dimensional scaling and correspondense analysis	1
Selection of variables in MR	1	Outlier detection and regression diagnostics	1
Computer Program Package	0.5	Use of Micro-computer and Large-scale computer (SAS)	1
Quantification Method I	1		
Exploratory Data Analysis	1		
Discussion	0.5	Discussion	0.5

* p denotes number of explanatory variables.

Specific curricula for the courses, "Design of Experiment" and "Multivariate Analysis", are presented in Tables 2 and 3. These educational programmes tend to illustrate various statistical techniques, which could be applied to respective occasions of industrial activities. It is, however, to be favored that those techniques are incorporated into a whole set of analyses of the existing procedures for locating ill-functioning processes. A following step is to modify or to relieve the failed processes, in which the statistical techniques could be an important tool for implementation.

Fig. 1. Problem-solving paths

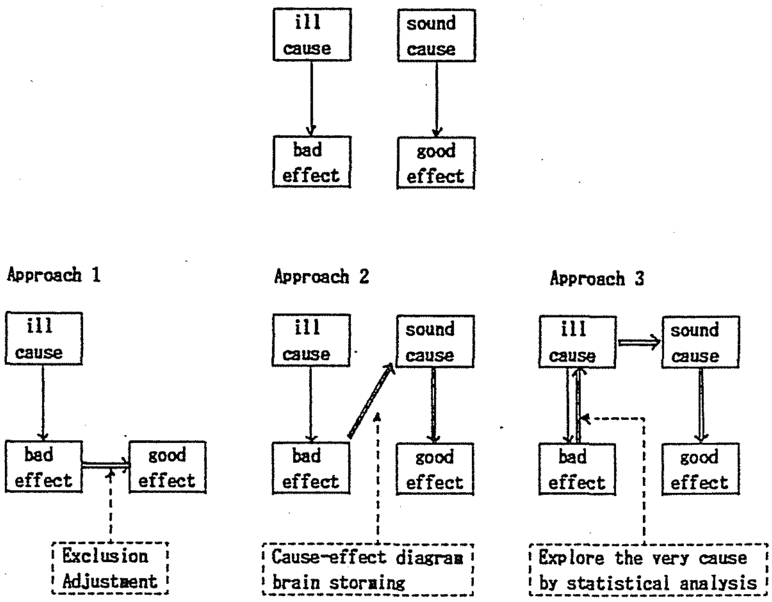


Fig. 1 will help to explain the problem-solving paths in the QC Education Scheme. In Approach 1 the goods of non-conformity (bad effect) resulting from a failed process (ill cause) are excluded or adjusted in an additional process to make satisfactory goods. The approach does not deal with the very cause so that the first-hand production of non-conformity goods continues. The cost of production cannot be reduced. In Approach 2 the cause-effect diagram is introduced to list plausible causes. Every measure put forth by all QC-circle members is undertaken for improving products. This approach works normally to the desired effect, but it provides no information on the whereabouts of failed processes. A series of excessive measures is very often taken up consequently, which leads to an added cost of products.

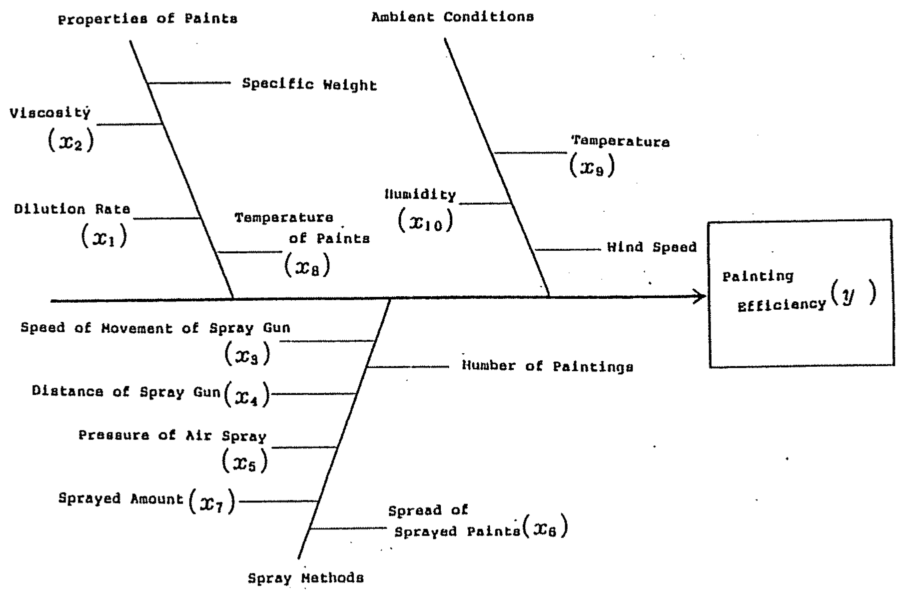
In Approach 3 every failed process can be located through an effective use of the multiple regression analysis on existing data and the design of experiment. This approach can indicate the processes which cause non-conformity goods. It will be able to lead to a complete elimination of failed processes and also to reduce non-conformity goods by a great degree.

3. Utilization of Multivariate Analysis Techniques

The author and five mechanical engineers have recently published a book entitled "Application of Multivariate Techniques in Automotive Industry". Those coauthors work for Toyota Motors and related companies. Eight cases from the production processes are dealt with in the publication. Referring to the book, the author illustrates examples of the utilization of multiple regression analysis and related techniques.

Case 1. Estimating optimum painting conditions of the body of a motor-car

Fig. 2. Cause-Effect Diagram for Painting Efficiency



The painting efficiency y (= amount of paints coated/amount of paints sprayed) is influenced by the quality and quantity of paints sprayed, spray methods and ambient conditions as shown in Fig. 2. We have found a model for estimating y by four variables – temperature of paints x_8, distance of spray gun x_4, spread of sprayed paints x_6 and sprayed amount x_7 – selected from 10 variables by using n = 38 observations. Table 4 shows several statistics for the best seven subsets with each specified number of variables. The "best" means the smallest values of Residual Sum of Squares (RSS) and the seven subsets are arranged in the increasing order of RSS values. Prediction Sum of Squares (PSS) and RSS are defined by

$$RSS = \Sigma_{i=1}^{n}(y_i - \hat{y}_i)^2, \qquad PSS = \Sigma_{i=1}^{n}(y_i - \hat{y}_i^*)^2,$$

where $\hat{y}_i$ and $\hat{y}_i^*$ (i=1,2,...,n) denote, respectively, the regression estimates derived from n observations and those from n-1 observations excluding the i-th observation. The author normally recommends adopting the subset with the minimum value of PSS.

Table 4. Statistics for the "best" Seven Subsets with Each Number of Variables

Number of Variables in Subsets	PSS	RSS	R**2	R'**2	R''**2	F-values for testing significance of the coefficients of selected variables									
						1	2	3	4	5	6	7	8	9	10
1	3103.85	2849.31	0.3775	0.3602	0.3438				10-						
	3523.68	3188.21	0.3034	0.2841	0.2657								10-		
	3934.33	3595.35	0.2145	0.1926	0.1719		9.8								
	4096.50	3720.46	0.1854	0.1620	0.1413									8.2	
	4644.17	4186.50	0.0853	0.0599	0.0358			3.4							
	4539.87	4213.77	0.0793	0.0538	0.0295										3.1
	4916.65	4342.85	0.0511	0.0248	-0.0002						1.9				
2	2443.75	2043.07	0.5536	0.5281	0.5039				10-		10-				
	2789.09	2411.21	0.4732	0.4431	0.4145				10-				6.4		
	3096.79	2640.71	0.4230	0.3901	0.3588		2.8		10-						
	3106.55	2713.76	0.4071	0.3732	0.3411			1.7	10-						
	3223.07	2773.00	0.3941	0.3595	0.3267				10-					1.0	
	3269.79	2787.43	0.3910	0.3562	0.3232						5.0		10-		
	3269.87	2808.16	0.3865	0.3514	0.3101			4.7					10-		
3	2035.45	1611.98	0.6478	0.6167	0.5872				10-		10-		9.1		
	2422.78	1932.13	0.5779	0.5406	0.5053		2.0		10-		10-				
	2433.85	1964.30	0.5708	0.5330	0.4970				10-		10-	1.4			
	2682.77	1967.97	0.5700	0.5321	0.4961				10-	1.3	10-				
	2514.65	1968.47	0.5699	0.5320	0.4960				10-		10-			1.3	
	2483.36	1980.14	0.5674	0.5292	0.4930			1.1	10-		10-				
	2620.55	2040.47	0.5542	0.5148	0.4775	0.0			10-		10-				
4	1924.93	1467.96	0.6793	0.6404	0.6035				10-		10-	3.2	10-		
	1987.03	1512.66	0.6695	0.6294	0.5914			2.2	10-		10-		10-		
	2084.12	1548.59	0.6617	0.6206	0.5817				10-		10-		8.9	1.4	
	2254.32	1558.12	0.6596	0.6183	0.5792				10-	1.1	10-		8.7		
	2205.84	1603.22	0.6497	0.6073	0.5670				10-		10-		9.0		0.2
	2172.59	1611.25	0.6480	0.6053	0.5648	0.0			10-		10-		8.8		
	2194.36	1611.87	0.6470	0.6051	0.5646		0.0		10-		10-		6.6		
5	2123.66	1417.62	0.6903	0.6419	0.5960				10-	1.1	10-	3.2	10-		
	1979.18	1421.20	0.6895	0.6410	0.5949				10-		10-	2.9	9.8	1.1	
	1964.88	1423.49	0.6890	0.6404	0.5943			1.0	10-		10-	2.0	10-		
	2020.76	1440.17	0.6853	0.6362	0.5895			2.4	10-		10-		10-	1.6	
	2014.43	1447.09	0.6838	0.6344	0.5876	0.6			10-		10-	3.6	10-		
	2036.95	1451.08	0.6830	0.6334	0.5874		0.4		10-		10-	3.5	6.8		
	2042.49	1456.46	0.6818	0.6321	0.5849				10-		10-	3.2	10-		0.3
6	2178.80	1345.83	0.7060	0.6490	0.5950				10-	1.7	10-	2.7	10-	1.7	
	2008.98	1368.00	0.7011	0.6433	0.5884			1.2	10-		10-	1.6	10-	1.3	
	2264.48	1373.71	0.6999	0.6418	0.5867			2.1	10-	1.6	10-		10-	2.2	
	2212.07	1382.22	0.6980	0.6396	0.5841			0.8	10-	0.9	10-	2.0	10-		
	2181.74	1384.64	0.6976	0.6389	0.5834		1.6		10-		10-	4.7	6.3		1.6
	2050.26	1390.56	0.6962	0.6374	0.5816	0.7			10-		10-	3.4	10-	1.3	
	2080.34	1398.70	0.6955	0.6366	0.5806			1.4	10-		10-	1.9	10-		0.7

The left half of Table 4 is for PSS, RSS and squared multiple correlation R^2, its adjusted value by degrees of freedom R'^2 and its doubly-adjusted value R''^2 derived from the expected value of PSS. The minimum value of PSS is shown 1924.93 in the subset with four variables mentioned above. In the right half of Table 4 are indicated the F-values for testing significance of the partial regression coefficients. 10- denotes that the F-value is larger than 10. The model becomes explicitly,

$$\hat{y} = 80.39 - 0.93x_4 - 4.38x_6 - 0.032x_7 - 0.45x_8, \quad R^2 = 06404.$$

The scatter diagram of 38 residuals $y_i - \hat{y}_i$ and regression estimates $\hat{y}_i$ reveals an outlier, which is excluded in the calculations to follow.

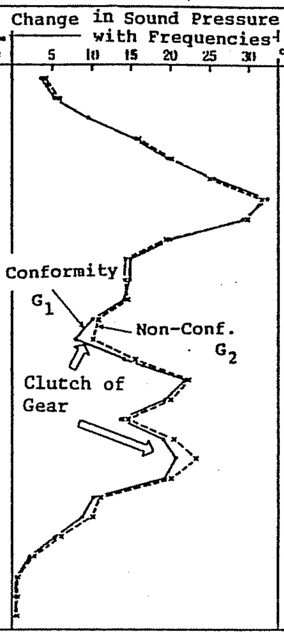
The similar analyses were also conducted on the other two criterion variables related to thickness of coating. It is found that the models for these two variables can be expressed by the same four explanatory variables as the painting efficiency.

Case 2. Analysis of gear noises of transmission

Sensory tests on gear noises of transmission by workers were conducted for discriminating two groups "Conformity and non-conformity". This sort of tests is laborious and the reproducibility is open to question. In this study, measurements of sound pressure at 28 frequencies were used for calculating stepwise discriminant functions and 5 frequencies at 40Hz, 160Hz, 630Hz, 2.5KHz and 10.0KHz were chosen as the most appropriate frequencies.

Fig. 3. Group Means of Sound Pressure at 28 Frequencies and Their Coefficients in the Discriminant Function

Variable	Frequency (Hz)	Group Means: Conformity G_1	Non-Conf. G_2	Difference $G_1 - G_2$	Coefficient[2]	t-value[3]
x_1	31.5	3.28	3.89	−0.61	−0.0012	0.20
O x_2	40	5.33	5.91	−0.58	0.0069	1.83
x_3	50	9.03	9.05	−0.02	−0.0032	0.62
x_4	63	16.07	16.39	−0.32	0.0014	0.19
x_5	80	19.07	19.77	−0.70	0.0003	0.14
x_6	100	26.21	25.95	0.26	−0.0028	0.58
x_7	125	32.36	32.11	0.25	−0.0063	1.24
O x_8	160	28.15	28.76	−0.61	−0.0178	3.41**
x_9	200	20.28	19.61	0.67	0.0017	0.22
x_{10}	250	14.89	14.56	0.33	0.0059	1.19
x_{11}	315	14.54	13.78	0.76	0.0004	0.07
x_{12}	400	13.15	14.29	−1.14	−0.0171	1.36
x_{13}	500	10.00	11.17	−1.17	0.0035	0.90
O x_{14}	630	7.74	10.68	−2.94	−0.0309	14.45**
x_{15}	800	14.40	16.87	−2.47	−0.0031	0.74
x_{16}	1.00k	22.41	22.58	−0.17	0.0061	1.01
x_{17}	1.25k	19.25	19.96	−0.71	0.0006	0.11
x_{18}	1.60k	13.34	14.41	−1.07	0.0015	0.23
x_{19}	2.00k	17.89	20.60	−2.71	−0.0009	0.16
O x_{20}	2.50k	21.46	23.53	−2.07	−0.0092	2.25*
x_{21}	3.15k	18.88	20.65	−1.77	−0.0069	1.38
x_{22}	4.00k	10.02	11.75	−1.73	−0.0022	0.45
x_{23}	5.00k	8.41	10.05	−1.64	0.0012	0.15
x_{24}	6.30k	5.47	6.86	−1.39	0.0008	0.15
x_{25}	8.00k	2.35	2.67	−0.32	0.0040	0.72
O x_{26}	10.00k	0.23	0.36	−0.13	−0.0316	2.46*
x_{27}	12.50k	0.52	0.52	0	0.0116	0.96
x_{28}	16.00k	0.30	0.45	−0.15	−0.0031	0.30



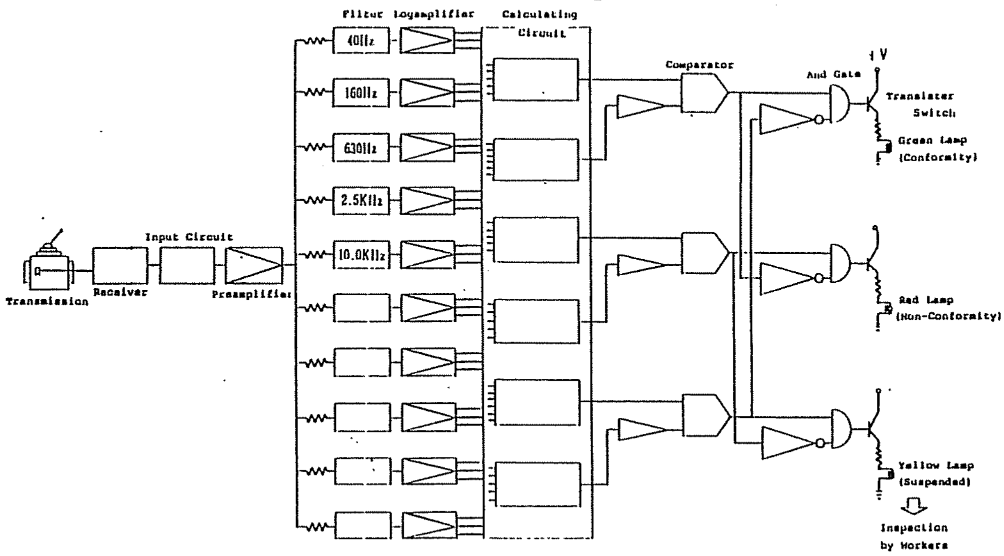
1) Mark o designates the variables selected by stepwise discriminant analysis

2) Coefficient of each variable in the discriminant function using 28 variables.

3) t-value is used for testing the significance of the coefficient, $t(71;0.05)=2.00$, $t(71;0.01)=2.66$

Fig. 3 shows the means of the two groups and their difference for each of 28 variables as well as the coefficients of the 28 variables in the discriminant function with corresponding t-values for testing their significance. It is interesting to note that at the above five frequencies both the differences of two-group means and the t-values are distinctly larger than those at other frequencies. This result has led to the design of an instrument for discrimination as shown in Fig. 4.

Fig. 4. An Instrument for Disciminating Gear Noises of Transmission



Case 3. Analysis of deficit in diecasting work

Table 5. Data with Category Numbers of 6 Items (factors)

No.	A	B	C	D	E	F	y
1	1	1	1	1	1	1	9.8
2	1	2	2	2	2	2	-1.9
3	1	1	3	3	3	1	2.7
4	1	1	1	3	2	1	4.2
5	1	2	2	3	3	2	-0.3
6	1	1	3	1	1	2	7.1
7	1	1	1	3	3	1	4.7
8	1	2	2	1	1	1	0.4
9	1	3	3	3	2	1	-2.4
10	1	1	1	1	1	1	13.5
11	1	2	2	2	2	2	-0.3
12	1	2	3	3	3	1	5.5
13	1	1	1	2	2	1	3.6
14	1	2	2	3	3	1	-0.1
15	1	2	3	1	1	2	11.0
16	1	1	1	3	3	2	1.2
17	1	2	2	1	1	1	0.6
18	1	3	3	2	2	1	6.9
19	2	1	3	2	1	1	-1.8
20	2	1	2	1	3	1	-3.7
21	2	3	1	3	2	2	-14.9
22	2	1	2	2	1	2	4.7
23	2	2	3	3	2	1	-5.2
24	2	3	1	1	3	1	-12.3
25	2	1	2	3	2	2	1.1
26	2	1	3	1	3	2	3.6
27	2	3	1	2	2	1	-5.8
28	2	1	2	1	3	1	0.6
29	2	1	2	1	3	1	2.4
30	2	2	3	2	1	1	-2.3
31	2	3	1	3	2	2	-10.7
32	2	1	2	2	1	2	-5.9
33	2	2	3	3	2	1	3.9
34	2	3	1	1	3	1	-3.0
35	2	1	2	3	2	1	9.8
36	2	2	3	1	3	2	-5.6
37	2	3	1	1	1	1	-10.3
38	3	1	3	1	2	2	-1.1
39	3	2	1	2	3	1	-3.6
40	3	3	2	3	1	1	-9.7
41	3	1	3	2	3	1	-2.2
42	3	2	1	3	1	2	1.1
43	3	3	2	1	2	1	-11.3
44	3	1	3	3	1	1	-4.6
45	3	2	1	1	2	1	-4.9
46	3	3	2	2	3	2	-5.0
47	3	1	3	2	2	2	14.6*
48	3	2	1	2	3	1	7.9
49	3	3	2	3	1	1	2.6
50	3	1	3	3	3	1	1.5
51	3	3	2	1	2	1	-1.1
52	3	1	3	3	1	1	14.2*
53	3	3	2	2	3	2	-2.4

Note: The bracketed observations have the same combination of categories. The differences of these data are used in estimating the variance of pure error.

The fractional-factorial design with 6 factors each at 3 levels was originally planned with the resulted specific weight data (coded) representing the degrees of deficit as shown in Table 5 but a few combinations of levels were proved to be difficult to realize. The resulting 53 combinations do not guarantee the orthogonality of the six factors. The comprehensive analysis of such complicated and incomplete data is conducted on computers by using "Quantification theory of the first type" of Hayashi, which can be considered as the "Multiple regression analysis for categorical data", or "Analysis of multi-way classification data with unequal numbers of observations in cells".

Table 6. An Illustration of Computer Output by Haga (Quantification Method of the First Type)

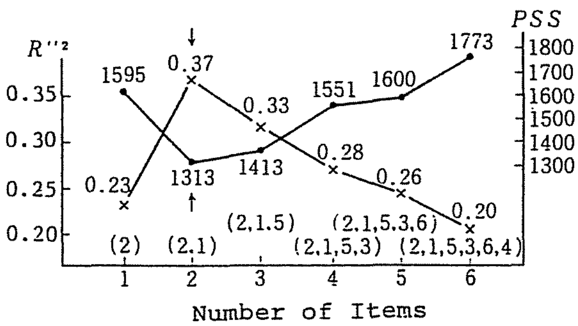
Item(factor) Category	No. of observations	Score b_{ij}	Corresponding t-value Differences in Scores			Range	Partial Correlation
1 (A)			1	2	3	6.141	0.464
1	18	3.185	0.000	3.241	1.623		
2	19	-2.956	6.141	0.000	-1.502		
3	16	-0.072	3.257	-2.884	0.000		
2 (B)			1	2	3	8.033	0.551
1	23	3.461	0.000	2.426	3.947		
2	16	-0.974	4.436	0.000	1.636		
3	14	-4.572	8.033	3.598	0.000		
3 (C)			1	2	3	2.077	0.176
1	16	-0.846	0.000	-0.205	-1.047		
2	19	-0.454	-0.392	0.000	-0.896		
3	18	1.231	-2.077	-1.686	0.000		
4 (D)			1	2	3	0.942	0.076
1	18	-0.001	0.000	-0.268	0.218		
2	15	0.539	-0.539	0.000	0.489		
3	20	-0.404	0.403	0.942	0.000		
5 (E)			1	2	3	1.882	0.160
1	16	0.878	0.000	0.302	0.998		
2	18	0.279	0.599	0.000	0.684		
3	19	-1.004	1.882	1.283	0.000		
6 (F)			1	2		0.753	0.072
1	35	0.255	0.000	0.455			
2	18	-0.498	0.753	0.000			

The computer output is illustrated in Table 6. The score b_{ij} represents the effect of the j-th category (level) of the i-th item (factor), subject to the condition:

$$\sum_j n_{ij} b_{ij} = 0 \qquad \text{for each } i,$$

where n_{ij} denotes the number of observations in the concerned category. In the middle three columns for the differences in the scores and their t-values, the lower off-diagonal elements for item A indicate the differences in scores like $b_{11} - b_{12} = 3.185 - (-2.956) = 6.141$. The corresponding t-values are given in the upper off-diagonal elements.

Fig. 5. Change of *PSS* and R''^2 with Number of Items



Selection of item in place of selection of each dummy variable is performed through the algorithm newly-developed by T. Haga of Science University of Tokyo. The progressive results in the intermediate steps are summarized in Fig. 5. The model accommodating only two items, temperature of melted aluminum (factor B) and speed of moulding (factor A), is proved to give the best subset by the criteria R'^2, R''^2 and PSS.

The above illustrations display clearly the characteristic points of our data analysis, which are summarized as follows:

1) Unbalanced data often derive from failed experimental designs. They can, however, be analysed by so-called "quantification methods", which are multiple regression and discriminant analyses for categorical data.

2) The model (or variable) selection in multiple regression analysis and detection of outliers are not connected normally with unique solutions. This is the characteristic of "Exploratory Data Analysis (EDA)".

3) The application of EDA will promote the development of SQC. The cases illustrated could serve as examples of methods to be adopted in such studies.

4. Concluding Remarks

The application of statistical methods should not be left exclusively for QC experts. Extension of the use of statistical techniques to those who concern the quality is required since data are acquired in their normal line of work. However, their understanding on the significance of data acquisition seems the least developed. Collecting data in the daily work could be decisive in reducing non-conformity goods. It is generally recommended that the collectors of data be supplied with 5W1H on their work: Who, When, Where, What, Why and How.

The educational programmes should include simple methods of graphical display and data analysis on the personal computer for general engineers. The advanced statistical analysis by use of program packages such as SAS on large-scale computers might be reserved in the educational programme for a fairly limited number of engineers.