Two Methods for Comparing Pareto Charts

RON S. KENETT

Tadiran, Ltd., Israel and State University of New York, Binghamton, New York 13901

Pareto Charts are a basic tool that derive their power from the Pareto principle. This work describes two statistical tests that complement the exploratory graphical analysis of the charts. The first test permits us to compare a given Pareto chart with another Pareto chart constructed over a different time period or process. The second test enables us to determine a relative ordering of sets of Pareto charts. The tests are presented in a step-by-step fashion with numerical examples.

Introduction

An article published in *Quality Progress* on the history of the "Pareto principle" (Juran [1975]) states: "Juran was (seemingly) the first to identify the phenomenon of the vital few and trivial many as a 'universal,' which seems applicable to many fields. Dr. J. Juran applied the name 'The Pareto Principle' to this universal. . . and applied the Lorenz curves to depict this universal in graphic form." These graphs were later labeled "Pareto Charts."

Pareto chart analysis is used to identify the major causes of phenomena like failures, defects, delays, et cetera. Typically Pareto charts complement control charts such as p-charts, np-charts, c-charts, or u-charts. They provide an analysis of what reasons were dominant in causing chronic or sporadic performance levels of a process. For a classic exposition of the subject, see Ishikawa (1976).

Pareto charts are also used to analyze data gathered in single time studies. For purposes of illustration we use an investigation of machine stoppages which greatly affected productivity (Price [1984]). The study was designed to find which of five possible reasons caused the machine to stop and for how long. Out of 26 stops that occurred in a period of 408 minutes the following distribution (by reasons labeled A-E) was observed:

A—13 stops (50%), B—5 stops (19%),

C—3 stops (11%), D—2 stops (8%),

E—3 stops (11%).

Dr. Kenett is Director of Statistical Methods at Tadiran, Ltd. and a Professor at the School of Management of the State University of New York, Binghamton. He is a Senior Member of ASQC.

A Pareto chart of this information is presented in Figure 1. Out of 108 minutes of total downtime the following distribution, by reasons, was observed:

A—22 min. (20%), B—26 min. (24%),

C—7 min. (6%), D—16 min. (15%),

E—37 min. (34%).

If frequency of stops is considered to be the principal nuisance, reason A is the obvious culprit. If, however, we look at improvement of overall efficiency, then eliminating reasons E and B will reduce downtime by 58%.

Quoting Price: "This is the usefulness of Pareto analysis; it signals those targets likely to yield maximum results by the deployment of limited effort. In acknowledging that there is little point in frittering away resources through fighting where the battle isn't raging, it pinpoints the most vulnerable areas of the enemy's line, so to speak. It is a technique which finds profitable employment when you are required to sort out a messy quality control situation. When customer's rejections are bombarding you so thick and so fast that you don't know where to begin, Pareto tells you." (Price [1984]). Pareto charts are similarly used in improvement projects and other "non fire-fighting activities."

In this paper we propose two statistical methods for comparing single and multiple Pareto charts. These methods identify differences in the relative weights of the reasons that contribute to a given phenomenon. The analysis of the mix of relative weights complements the investigation of number of occurrences or any other relative or absolute measures of the phenomenon's intensity. For example one might be interested in tracking the number of machine breakdowns and, in parallel, monitoring the distribution of

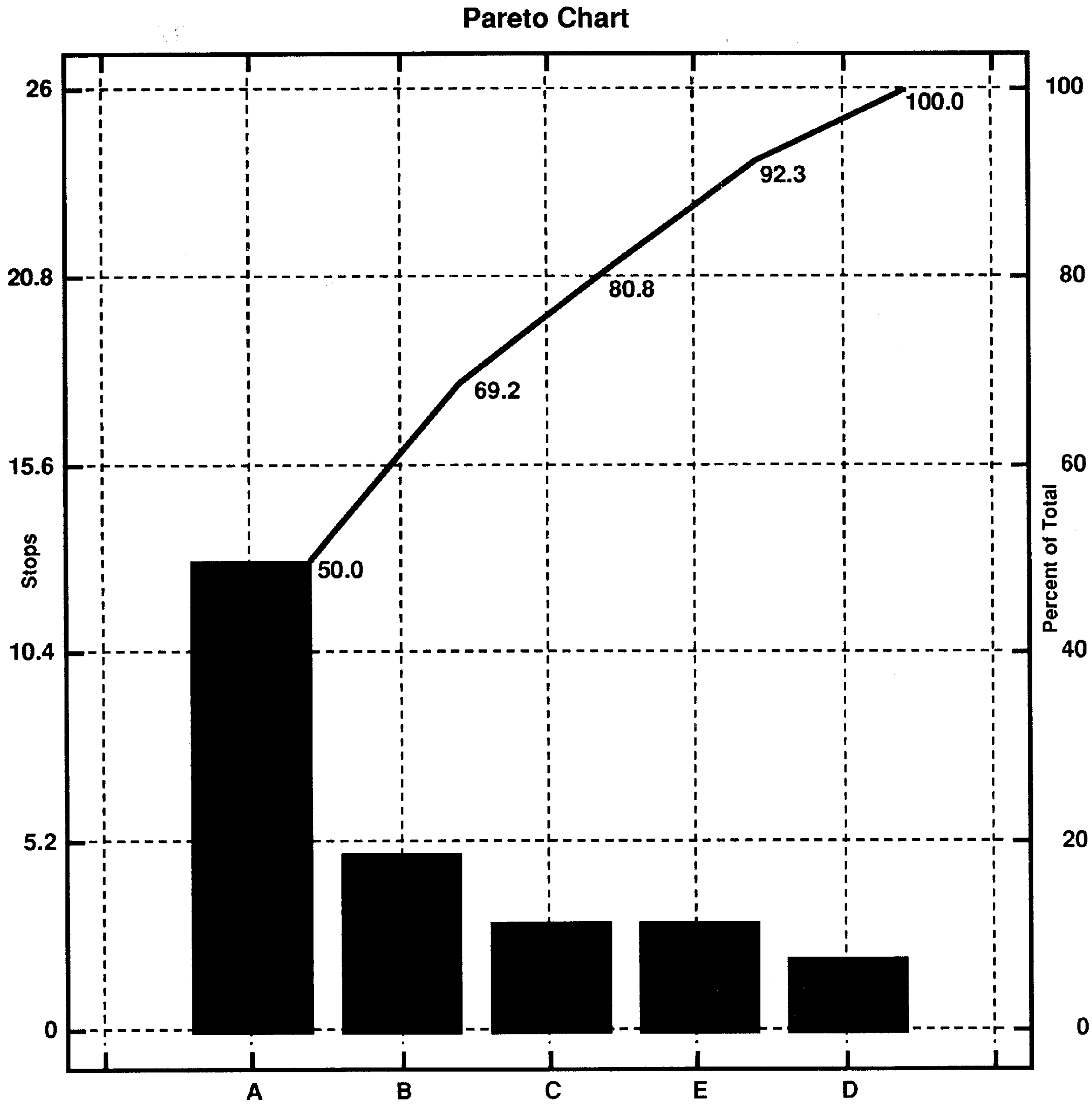

FIGURE 1. Pareto chart of reasons for machine stops.

breakdown reasons. We discuss applications of these methods to the detection of differences that can occur in time or across processes.

A Method for Comparing Two Pareto Charts

This section describes a method (called the M-test) for comparing two Pareto charts. The M-test has been shown (Fuchs and Kenett [1980]) to have larger asymptotic power than the standard chi-squared test in the presence of one or more outliers, for moderate and large numbers of categories. It is based on comparing the maximum of the absolute value of adjusted residuals against a certain critical value.

The M-test enables us to compare a given Pareto chart to a "standard" Pareto chart with the same defect categories. For example, the "standard" may be a Pareto chart consisting of historical defect causes and the given chart of interest may be a Pareto chart for the most current data. We assume that the percentages of the standard chart are exactly known and that the categories are mutually exclusive.

We demonstrate the procedure with the data presented in Figure 1 and the five arbitrary standards presented in Table 2. The M-test will determine if the data are consistent with any one of these five "standards," and, if significant differences are declared, point to the stoppage categories where the difference occurs.

Investigating differences between Pareto charts complements the analysis of overall process defect level behavior. Increases in defect levels may result from an increase across all categories which will not show in the Pareto chart comparison. On the other hand there may be no change in process defect levels but significant changes in the mix of the reasons causing the defects. This situation will not show on a control chart but will appear in a comparison of Pareto charts.

Assume the chart we want to compare to a "standard" Pareto chart is based on a total of N observations, subdivided into K categories, with n_i observations in category i. Let p_i be the expected proportion of observations in category i, according to the standard Pareto chart so that $n_1 + n_2 + \cdots + n_K = N$ and $p_1 + p_2 + \cdots + p_K = 1$.

The M-test consists of five steps:

1. Compute the expected number of observations in category i, E_i, by multiplying the total number of observations N by the standard expected percentage p_i of category i

$$E_i = Np_i, \quad i = 1, \cdots, K.$$

2. Compute the standard deviation S_i of the number of observations for category i using the standard category percentage

$$S_i = \sqrt{Np_i(1 - p_i)} \quad i = 1, \cdots, K.$$

3. Compute the adjusted residuals, Z_i, defined as

$$Z_i = \frac{n_i - E_i}{S_i} \quad i = 1, \cdots, K.$$

4. For a given significance level α and the given value of K, determine the critical value C of Z_i using Table 1 or the formula

$$C = \theta^{-1}(1 - \alpha/K)$$

where $\theta^{-1}\{Z\}$ is the inverse of the cumulative standard normal distribution evaluated at Z. The critical values thus derived are upper bounds and are not exact. In Fuchs and Kenett (1980) a lower bound is also provided. However since these bounds are tight the conservative approach taken here has a minimal effect on the procedure's significance level.

5. If all adjusted residuals Z_i are smaller, in absolute value, than C, no significant changes in the Pareto chart are declared. Cells with values of Z_i above C or below $-C$ are declared significantly different from the "standard," meaning a change in the distribution has occurred. If the change was unexpected an investigation should begin.

TABLE 1. Critical Values for Maximum Adjusted Residuals. The Table Shows Values of C for Various One Sided Significance Levels α and Number of Different Categories

K \ α	0.1	0.05	0.01
4	1.95	2.24	2.81
5	2.05	2.32	2.88
6	2.12	2.39	2.93
7	2.18	2.44	2.99
8	2.23	2.49	3.04
9	2.28	2.53	3.07
10	2.32	2.57	3.10
20	2.57	2.81	3.30
30	2.71	2.94	3.46

As an example, let us refer to the 26 machine stops discussed earlier. Table 2 shows the number of stops n_i by reason; the proportions p_i for five different "standard" distributions; the expected number of stops under these "standard" conditions E_i; and the corresponding adjusted residuals Z_i.

We want to determine, for illustration purposes, if the "standard" still holds or if our observations indicate a change in the distribution of reasons causing machine stops. From Table 1, with five categories ($K = 5$) the critical values are 2.05, 2.32, and 2.88 for 10%, 5%, and 1% one-sided significance levels, respectively.

The observations are consistent with standards 4 and 5 and no changes are declared there. In case 3 we have significant differences in cells A and E at the 10% level. In cases 1 and 2 the change can be determined at the 1% level. Thus, if case 1 were the "standard" Pareto chart, the distribution of the recent 26 stops shows that reason A is significantly more prevalent than in the standard and reason E is significantly less prevalent. If the standard represents common causes of the system, an investigation of these statistically significant differences is likely to show assignable causes.

TABLE 2. M-test Applied to Machine Stops Example (Bold Z_i Indicates 1% Significance)

		Total number of stops $N = 26$				
	Reason:	*A*	*B*	*C*	*D*	*E*
Standards	n_i	13	5	3	2	3
1.	p_i	0.1	0.2	0.1	0.1	0.5
	E_i	2.6	5.2	2.6	2.6	13.0
	Z_i	**6.8**	−0.1	0.3	−0.4	**−3.9**
2.	p_i	0.2	0.2	0.1	0.1	0.4
	E_i	5.2	5.2	2.6	2.6	10.4
	Z_i	**3.8**	−0.1	0.3	−0.4	**−3.0**
3.	p_i	0.3	0.2	0.1	0.1	0.3
	E_i	7.8	5.2	2.6	2.6	7.8
	Z_i	2.2	−0.1	0.3	−0.4	−2.1
4.	p_i	0.4	0.2	0.1	0.1	0.2
	E_i	10.4	5.2	2.6	2.6	5.2
	Z_i	1.0	−0.1	0.3	−0.4	−1.1
5.	p_i	0.5	0.2	0.1	0.1	0.1
	E_i	13.0	5.2	2.6	2.6	2.6
	Z_i	0.0	−0.1	0.3	−0.4	0.3

A Method for Comparing Multiple Pareto Charts

If there are several machines on the manufacturing floor, separate Pareto charts are typically constructed for each machine. Comparing individual machines by reasons of machine stops can be done using the *M*-test described in the previous section. If, however, we are interested in overall differences (e.g. over time), a different approach is needed. We propose to compare separate Pareto charts obtained every time period with one "standard" set of charts. This global comparison will tell what time periods most resembled the standard set and what time periods were different. We will use the sign test, which is based on relative comparisons and is not sensitive to the actual values of the numbers compared. Such a non-parametric procedure can lead to "statistical significance" and conceal the "practical significance" of the test outcome. Careful interpretation of the results is therefore necessary.

The proposed method consists of three steps:

1. Determine a set of "standard" Pareto charts, each based on the same K categories.
2. For each time period and for each machine compute the chi-squared distance of the observations from the "standard."

$$\text{chi-squared distance} = \sum_{i=1}^{K} \frac{(n_i - E_i)^2}{E_i}$$

where E_i is determined as in step 1 of the *M*-test.

3. Compare any two time periods relative to the standard, say period 1 and period 2, by determining the number of machines with larger chi-squared distances in period 1 than in period 2. Significant differences are to be determined using the sign test.

In a set consisting of six Pareto charts, the most extreme comparison is 6:0. This will happen if the six chi-squared distances in period 1 are all greater than those in period 2. A comparison of 5:1 or 6:0 has an 11% significance level for declaring that the charts of period 1 are more different from the standard charts than those of period 2. Ratios of 4:2 or 3:3 indicate no such differences. In fact we can create a partial ordering of weeks based on their relative distances from the standard set. (For further details on the method and other applications see Kenett [1983]).

For purposes of illustration we use the data in Table 3. The chi-squared distances are presented in Table 4. We see that chi-squared values in period 1 are larger than those of period 2, for machines 2, 3 and 5. This results in a 3:3 comparison. We therefore declare that overall period 1 was no more different from the standard than period 2. If a new maintenance program had been initiated, say in period 2, we would declare it to have had no significant effect on machine performance.

Schematically we can show this conclusion as

period 1

Standard |——————————————|

period 2

Summary and Conclusions

Pareto charts are a basic tool for the analysis of categorical data. Visual inspection of the chart is a first step in the data analysis and the Pareto principle typically provides much of the insight. Additional analyses can yield more knowledge. In particular the incorporation of statistical tests to determine the randomness due to "common" causes and that due to "assignable" causes will help prevent what W. E. Deming calls "going down the milky way." The two methods we described in this paper address two different issues. For comparing a given Pareto chart with a separate chart constructed over a period of time, we propose the *M*-test. When interested in a relative ordering of sets of Pareto charts we suggest a technique based on chi-squared distances and the sign test. These tests add a new dimension to the analysis of Pareto charts.

TABLE 3. Frequency of Machine Stops by Reason over Two Time Periods and Stop Category Percentages under the "Standard" Condition

	Machine					
Reason	*M1*	*M2*	*M3*	*M4*	*M5*	*M6*
			(a) Period 1			
A	18	25	14	5	9	3
B	16	12	9	2	6	4
C	5	4	3	7	6	9
D	2	4	2	2	2	2
E	4	5	6	10	9	6
Total	45	50	34	26	32	24
			(b) Period 2			
A	26	32	28	6	21	3
B	13	7	3	4	5	5
C	2	6	4	2	7	3
D	2	4	4	1	3	1
E	4	12	10	5	7	4
Total	47	61	49	18	43	16
			(c) Standard			
A	0.48	0.51	0.51	0.25	0.40	0.15
B	0.31	0.17	0.14	0.14	0.15	0.23
C	0.08	0.09	0.09	0.20	0.17	0.30
D	0.04	0.08	0.07	0.07	0.07	0.08
E	0.09	0.15	0.19	0.34	0.21	0.24
Total	1.00	1.00	1.00	1.00	1.00	1.00

TABLE 4. Chi-squared Values of Machine Stops over Two Time Periods

	Machine					
Time Period	*M1*	*M2*	*M3*	*M4*	*M5*	*M6*
Period 1	1.34	2.41	4.54	1.88	2.28	0.97
Period 2	1.64	2.07	2.68	2.34	1.63	1.36

Acknowledgments

The constructive comments of the Editor, Drs. C. Fuchs and D. Steinberg, and two anonymous referees helped improve this article and are gratefully acknowledged.

References

FUCHS, C. and KENETT, R. S. (1980). "A Test for Detecting Outlying Cells in the Multinomial Distribution and Two-Way Contingency Tables." *Journal of the American Statistical Association* 75, pp. 395–398.

ISHIKAWA, K. (1976). *Guide to Quality Control,* Revised ed. Asian Productivity Organization, Tokyo, Japan.

JURAN, J. M. (1975). "The Non-Pareto Principle; Mea Culpa." *Quality Progress* 8, pp. 8–9.

KENETT, R. S. (1983). "On an Exploratory Analysis of Contingency Tables." *The Statistician* 32, pp. 395–403.

PRICE, F. (1984). *Right First Time: Using Quality Control for Profit,* Gower Publishing Company, Brookfield, VT.

Key Words: *Pareto Chart, Quality Management, SPC, SQC.*