

Guide to quality control and performance improvement using qualitative (attribute) data —

Part 3: Technical aspects of attribute charting: special situation handling

ICS 03.120.30

Confirmed November 2008

Committees responsible for this British Standard

The preparation of this British Standard was entrusted to Technical Committee SS/4, Statistical process control, upon which the following bodies were represented:

- Association for Road Traffic Safety and Management (ARTSM)
- BAE Systems
- British Standards Society (BSS)
- Clay Pipe Development Association (CPDA)
- Federation of Small Businesses (FSB)
- Gauge and Tool Makers Association (GTMA)
- General Domestic Appliances Ltd.
- Institute of Quality Assurance
- National Physical Laboratory
- Royal Statistical Society

This British Standard was published under the authority of the Standards Policy and Strategy Committee on 31 October 2003

© BSI 31 October 2003

First published as BS 5701 February 1980

Amendments issued since publication

Amd. No.	Date	Comments

The following BSI references relate to the work on this British Standard:
Committee reference SS/4
Draft for comment 02/400887 DC

ISBN 0 580 42736 6

Contents

	Page
Committees responsible	Inside front cover
Foreword	ii
1 Scope	1
2 Normative references	1
3 Terms, definitions and symbols	1
4 Qualitative data fundamentals	1
5 Limitations of standard control chart limits: appropriate counter-measures	3
6 Specialized design control charts	18
7 Attribute charts for measured data situations	22
Annex A (normative) Poisson and binomial tests	26
Annex B (normative) Poisson-based control chart limits	28
Bibliography	31
Figure 1 — Illustration of the variation in distribution shape with changes in the mean	5
Figure 2 — Count data control limit selection flow chart	6
Figure 3 — “c” chart for accidents	7
Figure 4 — “c” chart for orders per day	9
Figure 5 — Histogram of adjustment data	10
Figure 6 — Normal probability plot of adjustment data	10
Figure 7 — Control chart for number of adjustments per unit	11
Figure 8 — Histogram of fabric fault data	12
Figure 9 — Run chart of fabric faults indicating significant improvement in performance	12
Figure 10 — Control chart for number of faults per roll	13
Figure 11 — Change in shape of the binomial distribution with different means	14
Figure 12 — Classified data distribution selection flow chart	15
Figure 13 — Control chart for non-conforming weld data	18
Figure 14 — Universal “u” chart for stud threads	20
Figure 15 — How to select the appropriate control chart for measurable data	22
Figure 16a) — Two-way attribute control chart	23
Figure 16b) — Illustration of process diagnosis using two-way control charts	24
Table 1 — Effect of method of calculation of action control limits in terms of value of the mean	4
Table 2 — Example of difference between conventional and Poisson-based control limits	7
Table 3 — Conventional v Poisson control limits	8
Table 4 — Binomial probabilities for $n = 500$; $p = 0.019\ 8$	17
Table 5 — Plotting data for standardized control charts	18
Table 6 — Example of tabulation for setting up a universal “u” chart	19
Table 7 — Formulae for setting up a demerits attribute chart	20
Table 8 — Results of twenty consecutive audits in multiple characteristic control chart format	21
Table A.1 — Critical values of variance ratio for testing binomial or Poisson distribution assumptions	27
Table B.1 — Poisson-based upper action control limits (set at 0.001 35 probability)	28
Table B.2 — Poisson-based lower action control limits (set at 0.001 35 probability)	29
Table B.3 — Poisson-based upper warning control limits (set at 0.022 8 probability)	29
Table B.4 — Poisson-based lower warning limits (set at 0.022 8 probability)	29

Foreword

BS 5701-3:2003 partially supersedes BS 5701:1980 and BS 2564:1955 and all four parts of BS 5701 together supersede BS 5701:1980 and BS 2564:1955, which are withdrawn.

Qualitative data can range from overall business figures such as percentage profit to detailed operational data, such as percentage absenteeism, individual process parameters and product/service features. The data can either be expressed sequentially in yes/no, good/bad, present/absent, success/failure format, or as summary measures (e.g. counts of events and proportions). For measured data control charting refer to BS 5702-1. The focus throughout the BS 5701 family of standards is on the application of attribute control charts to monitoring, control and improvement. The roles of associated, mainly pictorial, diagnostic, presentation and performance improvement tools, such as priority (Pareto) diagrams, cause and effect diagrams and flow charts are also indicated.

This aim of BS 5701 is to be readily comprehensible to the very extensive range of prospective users and so facilitate widespread communication, and understanding, of the method. As such, it focuses to the greatest extent possible on a practical non-statistical treatment of the gathering and charting of qualitative data.

BS 5701-1 demonstrates the business benefits, and the versatility and usefulness of a very simple, yet powerful, pictorial control chart method for monitoring and interpreting qualitative data. In BS 5701-1, the treatment of charting of qualitative data is essentially at appreciation level. However, it is intended to provide adequate information for a gainful first application, by a typical less statistically inclined user, in many everyday situations.

BS 5701-2 continues to focus on the application of standard attribute charts to the monitoring, control and improvement of business processes. This is done at a technical level more suitable for practitioners.

BS 5701-3 concentrates on the statistical basis of, and technical rationale for, attribute control charting. It also gives guidance on dealing with special situations.

BS 5701-4 deals with measuring and improving the quality of decision making in the classification process itself.

A British Standard does not purport to include all the necessary provisions of a contract. Users of British Standards are responsible for their correct application.

Compliance with a British Standard does not of itself confer immunity from legal obligations.

Summary of pages

This document comprises a front cover, an inside front cover, pages i and ii, pages 1 to 31 and a back cover.

The BSI copyright notice displayed in this document indicates when the document was last issued.

1 Scope

BS 5701-3 describes the technical and statistical foundations of attribute control charting. It also gives guidance on dealing with special situations.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

BS EN ISO 9000:2000, Quality management systems — Fundamentals and vocabulary.

BS ISO 3534-1, Statistics — Vocabulary and symbols — Part 1: Probability and general statistical terms.

BS ISO 3534-2, Statistics — Vocabulary and symbols — Part 2: Applied statistics.

3 Terms, definitions and symbols

For the purposes of this part of BS 5701, the terms, definitions and symbols given in BS ISO 3534-1, BS ISO 3534-2 and BS EN ISO 9000:2000, Clause 3 apply with one exception.

The exception is intended for purposes of simplicity. It relates to the symbols for population parameters, sample statistics and their realized values. It is standard statistical practice for population parameters to be symbolized by lower case Greek letters in italics. For example, mean and standard deviation are denoted by μ and σ . A population parameter is a summary measure of the value of some characteristic of a population.

A sample statistic is a summary measure of some observed value of a sample. Sample statistics are symbolized by upper case Latin letters in italics. For example, mean and standard deviation are denoted by $\bar{X}$ and S .

Realized values of sample statistics to be symbolized by lower case Latin letters in italics. For example, mean and deviation are denoted by $\bar{x}$ and s .

Upper case Latin letters in italics are used in Annex A and its applications.

4 Qualitative data fundamentals

4.1 General

Two general classifications of data are quantitative data and qualitative data. This standard focuses on qualitative data. Qualitative data are divided for convenience, into two categories, *classified data* and *count data*.

4.2 Classified data and the binomial distribution

With classified data, each item of data is classified as being one of a number of categories. Frequently the number of categories is two, i.e. a binary situation where, for instance, results are usually expressed as 0 and 1, or as, good/bad, success/failure, profit/loss, in/out, or presence/absence of a particular characteristic.

Data having two classes is termed “binomial” (binomial = “two names”) data. A measure can be inherently binomial, e.g. where a profit or loss is made, or of someone is in or out. Sometimes it is arrived at indirectly by categorizing some other numerical measure. Take, for instance, a case where telephone calls are classified on whether or not they last more than 10 min or, perhaps, whether or not they are answered within six rings.

The binomial distribution is a statistical distribution giving the probability of obtaining a specified number of particular outcomes (e.g. successes) in a specified number of independent trials with a constant probability of success in each.

The conditions that need to be satisfied for a binomial distribution are:

- a) there is a fixed number of trials, n ;
- b) only two possible outcomes are possible at each trial;
- c) the trials are independent;
- d) there is a constant probability of a particular outcome e.g. success, “ p ”, in each trial;
- e) the variable is the total number of successes in n trials.

$$\text{mean} = np \quad (1)$$

$$\text{standard deviation} = \sqrt{np(1-p)} \quad (2)$$

The binomial distribution is very cumbersome to calculate. When dealing with control limits other distributions can be used as approximations to the binomial. For example, if p is small, in Equation 2, $(1-p)$ approaches 1 and the standard deviation of the binomial distribution approximates to the square root of its mean. Equations (1) and (2) then become identical to equations (3) and (4) of the Poisson distribution.

In standard statistical process control methodology, binomial data is monitored using:

- i) “ p ” charts for proportions, particularly when the sample size is variable; and
- ii) “ np ” charts for numbers from samples of constant size.

4.3 Count data and the Poisson distribution

Count data relates to counts of events where each item of data is the count of the number of particular events per given period of time or quantity of product. Instances are: number of accidents or absentees per month, number of operations or sorties per day, number of incoming telephone calls per minute, or number of non-conformities per unit or batch.

The Poisson distribution has two principal parts to play in this standard:

- i) as an approximation to the more cumbersome binomial when:

n is large and p is small, say, $n > 20$ and $p \leq 0.1$ (i.e. $\leq 10\%$);

- ii) as a distribution in its own right. The binomial distribution applies to the situation where an item, say, is classified as conforming or non-conforming. Here one can count either the number of items non-conforming or the number of items conforming. Poisson data arises when one can count only, say, the number of faults in an item not the number of non-faults. The conditions that need to be satisfied for a Poisson distribution are that events occur randomly, singly, uniformly and independently.

The validity of the Poisson model hinges on the independence of events and their occurrence at an average rate that is assumed to be stable (in the absence of special causes).

Two fundamental features of the Poisson distribution are:

- a) the connection between its standard deviation and mean; and
- b) the results of summation of two independent Poisson variables.

Regarding a), the mean and standard deviation of a Poisson distribution are given by:

$$\text{mean} = m \quad (3)$$

$$\text{standard deviation} = \sqrt{m} \quad (4)$$

This relationship is useful for establishing the validity of the Poisson model for a particular set of data. Whilst many real life situations can suggest that Poisson is a reasonable underlying model it would be unwise to take this for granted. Sometimes a degree of clustering of events is present, for example, of accidents and faults. This has the effect of increasing the spread beyond that of the Poisson distribution. This wider spread, termed over-dispersion, can give rise to a higher false alarm rate when monitoring using control charts.

To reduce the risk of assuming an incorrect model a simple “dispersion” test is often recommended (see Annex A). This involves comparing the sample variance of the data with the mean.

Equation (4) indicates that the variance (the square of the standard variation) is equal to the mean for a Poisson distribution. Hence if:

- a) sample variance equals the mean, the Poisson model is plausible;
- b) sample variance is “much greater or smaller than” the mean, the Poisson model is implausible.

This variance comparison can be supported by other methods such as a graphical evaluation of the data and by goodness of fit tests such as the chi-squared test for differences between expected and observed frequencies.

Regarding b), distribution of the sum of independent Poisson variables is still Poisson. For example:

“if two sets of Poisson data are independent with separate means m_1 and m_2 , their sum is also Poisson with mean, $(m_1 + m_2)$ ”.

This fact has considerable impact on counts where multiple characteristics are involved. Instances are adding together different types of flaws, absenteeism for different reasons and injuries of different kinds. Another application is where events of a particular kind have a very small mean so that counts are frequently zero. This enables one to increase the number of sampled items (e.g. counts per month rather than per day or week).

In standard statistical process control, Poisson data is monitored using:

- 1) “c” charts for numbers (counts of events) from samples of constant size; and
- 2) “u” charts for proportions (number of events per sample), particularly when the sample size is variable.

5 Limitations of standard control chart limits: appropriate counter-measures

5.1 Bases for selection of control chart limits for count data

5.1.1 Introduction

It is common practice throughout the world to base standard control chart limits on the assumption of normality of data. Control limits are then calculated thus:

- a) action control limits = mean $\pm$ 3 standard deviations, of the plotted statistic;
- b) warning control limits = mean $\pm$ 2 standard deviations, of the plotted statistic.

In terms of count data using a “c” chart, these equations become:

$$\text{action control limits} = \bar{c} \pm 3\sqrt{\bar{c}} \quad (3)$$

$$\text{warning control limits} = \bar{c} \pm 2\sqrt{\bar{c}} \quad (4)$$

However, it should be appreciated that the “normal” distribution is symmetrical and bell shaped whereas the actual “Poisson” distribution, representative of count data, is frequently not so.

Does this really matter?

There are those who would argue that it doesn’t really matter very much, as control chart limits are based on quite arbitrary considerations anyway. Anyhow, even if it does matter somewhat, taking it into account would make things more difficult and is likely to put a lot of people off using a very simple and useful tool.

Others would argue that it is purely a matter of the method of establishing control limits and does not add any complexity to the application of a control chart by the users. To ensure the validity, credibility and maximum effectiveness of control charts, there is a need to know both the consequences of the choice of control limits, and the appropriate countermeasures that can be taken, by people technically responsible for them.

This standard takes the view that people technically responsible for control charts need to be aware of the consequences of, and the measures required to correct for, any errors arising from the use of standard control chart limits for count data. This enables judgements to be made on the proper course of action to take in particular circumstances.

5.1.2 Consequences

Figure 1 compares the shape of the “Poisson” distribution for various values of the data mean.

The figure shows that standard conventional control limits, based on symmetry about the mean, are sometimes not a good reflection of the actual situation. Both the upper and lower control limits are adversely affected. The consequences of this are manifold. With respect to the application of the appropriate Poisson distribution, the symmetrical distribution used for standard conventional control chart limits, the most generally used, gives:

- a) tighter upper control limits, progressively so with lower values of the mean;
- b) slacker lower control limits, progressively so with lower values of the mean;
- c) impossible negative values for low means.

An example of a situation where standard limits are not effective is the calibration of a particular meter that is prone to “sticking” to give the occasional zero reading. It is looked upon, erroneously, as “in control” in terms of the standard lower control limit, as zero is assumed because it calculates as negative with the low values of the mean experienced. Use of the Poisson-based control limit method provides correct lower control limits to indicate an “out of control” situation.

These differences in Poisson distribution patterns affect the risk probabilities associated with control limits at the mean ± 2 standard deviations. When the distribution is not near normal, as is frequently the case with attribute data, these probabilities will vary from those calculated using standard conventional practice based on normality of data. For different values of process mean the risk associated with a decision concerning control of the process will change.

Constancy of risk, at the level obtained with standard Shewhart-type control limits for measured data, can usually readily be achieved using Poisson-based control limits rather than the conventional limits more generally used. These Poisson-based control limits are set out for various means in Annex B.

Figure 1 shows that for:

- a) mean = 1, the pattern is decidedly non-symmetrical;
- b) mean = 5, there is a tendency towards skewness to the right;
- c) mean = 10, the skewness is less pronounced than at a mean = 5;
- d) mean = 20, a near symmetrical normal pattern emerges.

The variation in the value of control chart limits between the conventional and Poisson methods of calculation (see Annex B) is shown in Table 1.

Table 1 confirms that the use of conventional control limits, rather than the more appropriate Poisson limits, can give rise to a fair amount of error and, in fact, in some cases, give impossible negative values.

Table 1 — Effect of method of calculation of action control limits in terms of value of the mean

Type of control limit	Mean = $\bar{c}$	Control limit values	
		Conventional $\bar{c} \pm 3\sqrt{\bar{c}}$	Poisson
Upper action	1	4.0	5.7
	5	11.7	13.7
	10	19.5	21.7
	20	33.4	35.7
Lower action	1	negative	0
	5	negative	0
	10	0.5	1.3
	20	6.6	7.3
NOTE Conventional action control limits are based on the standard formula $\bar{c} \pm 3\sqrt{\bar{c}}$			

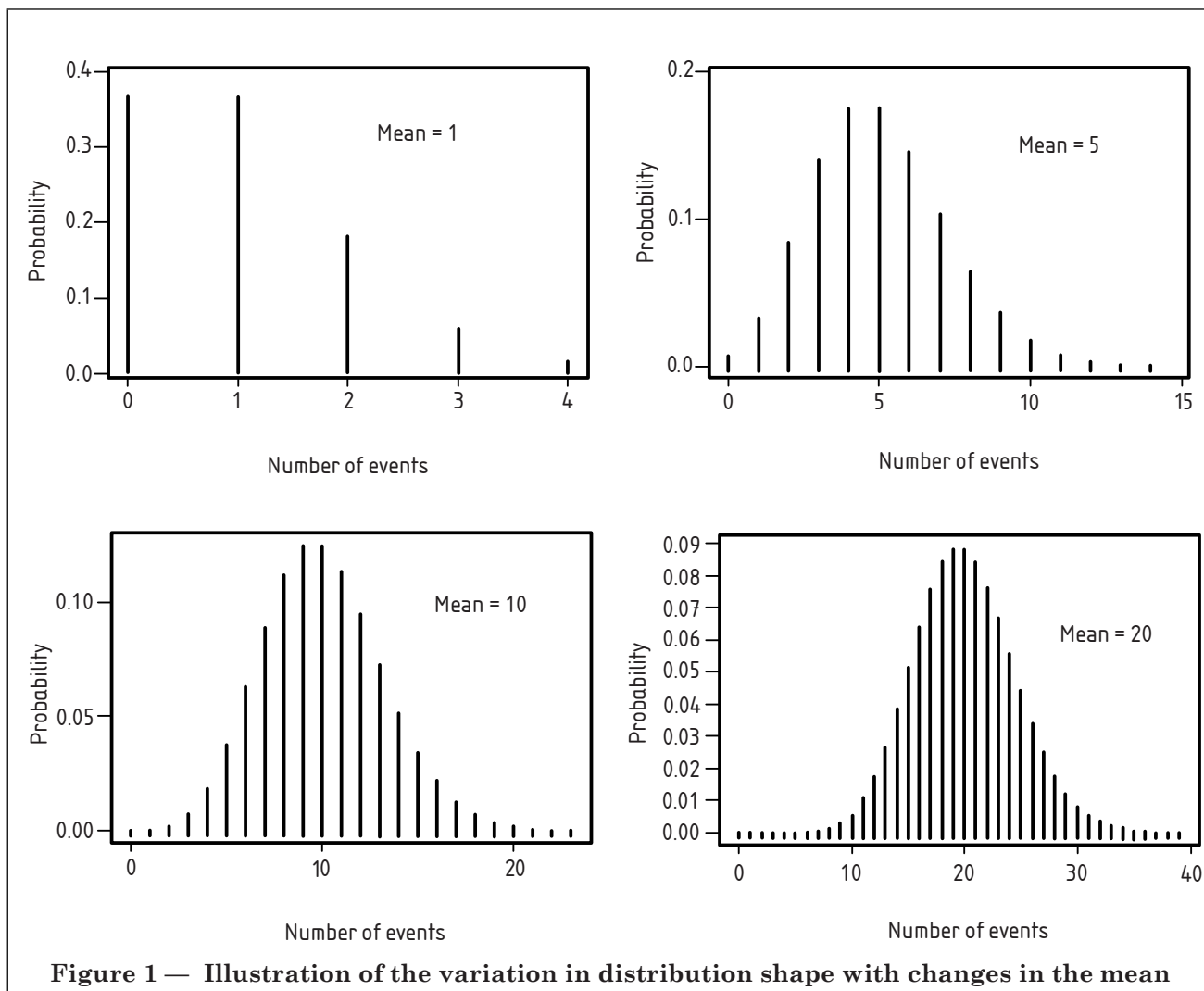
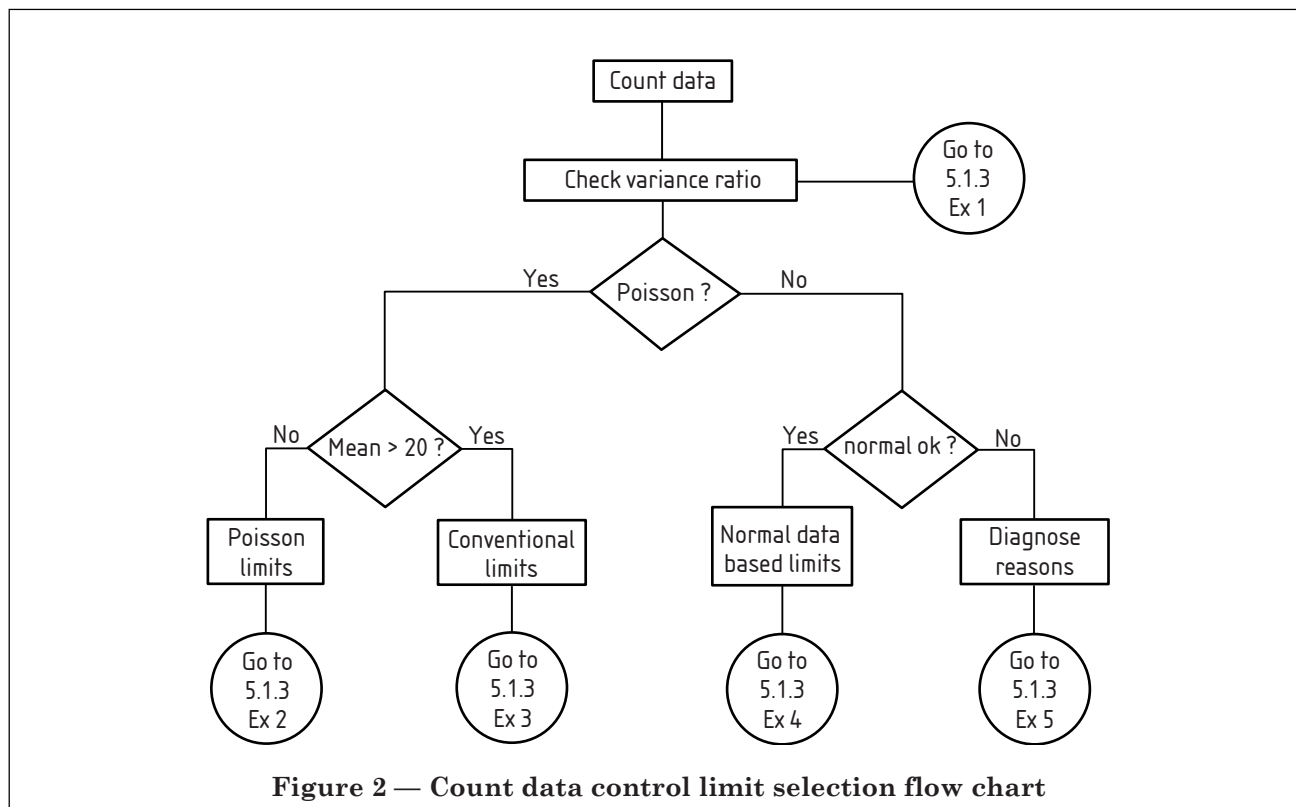


Figure 1 — Illustration of the variation in distribution shape with changes in the mean

5.1.3 Countermeasures

The rational approach to the setting of control limits for count data is indicated in the flow chart of Figure 2.



The flow chart of Figure 2 is used to construct data-based control chart limits for count data. Stages are as follows.

- 1) Apply Poisson variance ratio test to the count data. Use Annex A: Steps 1 to 3. (See Example 1.)
- 2) If the test indicates that the data:
 - a) appears to confirm a Poisson distribution go to Stage 3;
 - b) if not go to Stage 5.
- 3) If the data mean is less than 20, select Poisson-based control limits using Annex B (see Example 2) as conventional control limits give rise to a margin of error.
- 4) If the data mean is of the order of 20 or more, an appropriate approximation is to select conventional control limits using standard formulae. (See Example 3.)
- 5) Compare the data distribution with the normal distribution.
 - a) If the comparison is good select normal data-based control limits, using the observed rather than theoretical Poisson-based standard deviation. It is advisable also to investigate the possible causes for the unexpected non-Poisson type distribution using the methods of 5b). (See example 4.)
 - b) If not, diagnose the reasons for abnormality using Annex A: Steps 4 and 5. (See Example 5.)

EXAMPLE 1 Poisson variance ratio test

Project

The number of accidents, per four-week period, in a particular organization, are shown for 25 consecutive time periods.

12 5 7 10 9 5 11 6 7 9 5 11 9 4 12 6 11 9 7 4 11 10 8 7 6

Apply the Poisson variance ratio test as shown in Annex A.

Solution

Observed mean of data = $\bar{X} = \frac{1}{g} \sum X = 8.04$ (or use a calculator) (where g = number of samples)

Observed variance of data = $\frac{1}{g-1} \left[\sum (X - \bar{X})^2 \right] = 6.46$ (or use a calculator and square the readout from the σ_{n-1} key).

Estimated theoretical variance of data (based on Poisson distribution) = sample mean = 8.04

$$V = \frac{\text{observed variance}}{\text{estimated theoretical variance}} = \frac{S^2}{\bar{X}} = \frac{6.46}{8.04} = 0.80$$

From Table A.1, for number of samples = $g = 25$, critical values are: 0.41 to 1.90 at the 1 % level of significance. $V = 0.80$ is within this range therefore there is no reason to suppose the distribution of the data is other than Poisson.

EXAMPLE 2 Poisson-based control limits*Project*

Establish a control chart, with action limits, for the data of Example 1.

Solution

The solution to Example 1 indicates that there is no reason to suppose the distribution of the data is other than Poisson. A run chart of the data indicates that the data appears reasonably stable.

The mean = 8.04. This value is less than 20 therefore the Poisson distribution is used to calculate control limits. Reading from the tables in Annex B, the appropriate Poisson control limits are as shown in Table 2. The corresponding control limits for the inappropriate conventional method are also given for comparison to indicate the degree of error involved.

Table 2 — Example of difference between conventional and Poisson-based control limits

Type of control limit	Control limit values for mean = 8.04	
	Poisson-based (appropriate)	Conventional (appropriate) ^a
Upper action	18.7	16.5
Lower action	0.3	Negative
Upper warning	14.7	13.7
Lower warning	2.3	2.4

^a calculated from the conventional formulae:
 action control limits = mean $\pm 3\sqrt{\text{mean}}$, of the plotted statistic = $\bar{c} \pm 3\sqrt{\bar{c}}$
 warning control limits = mean $\pm 2\sqrt{\text{mean}}$, of the plotted statistic = $\bar{c} \pm 2\sqrt{\bar{c}}$

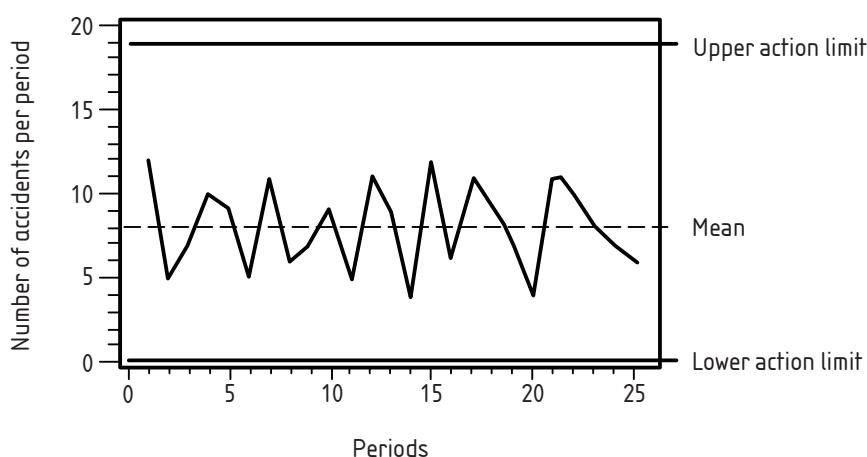


Figure 3 — “c” chart for accidents