PD ISO/TR 16705:2016

BSI Standards Publication

Statistical methods for implementation of Six Sigma — Selected illustrations of contingency table analysis

bsi.

National foreword

This Published Document is the UK implementation of ISO/TR 16705:2016.

The UK participation in its preparation was entrusted to Technical Committee MS/6, Methodologies for business process improvement using statistical methods.

A list of organizations represented on this committee can be obtained on request to its secretary.

This publication does not purport to include all the necessary provisions of a contract. Users are responsible for its correct application.

ISBN 978 0 580 74469 3

ICS 03.120.30

Compliance with a British Standard cannot confer immunity from legal obligations.

This Published Document was published under the authority of the Standards Policy and Strategy Committee on 31 August 2016.

Amendments issued since publication

Date	Text affected

TECHNICAL REPORT

ISO/TR 16705

First edition
2016-08-15

Statistical methods for implementation of Six Sigma — Selected illustrations of contingency table analysis

Méthodes statistiques pour l'implémentation de Six Sigma — Exemples sélectionnés d'application de l'analyse de tableau de contingence

Reference number
ISO/TR 16705:2016(E)

Contents

Page

Foreword ..... **iv**
Introduction ..... **v**
1 Scope **1**
2 Normative references **1**
3 Terms and definitions **1**
4 Symbols and abbreviated terms **2**
5 General description of contingency table analysis **2**
5.1 Overview of the structure of contingency table analysis 2
5.2 Overall objectives of contingency table analysis 3
5.3 List attributes of interest 3
5.4 State a null hypothesis 3
5.5 Sampling plan 3
5.6 Process and analyse data 4
5.6.1 Chi-squared test 4
5.6.2 Linear trend test 6
5.6.3 Correspondence analysis 6
5.7 Conclusions 7
6 Description of Annexes A through D **7**
Annex A (informative) **Distribution of number of technical issues found after product release to the field** **8**
Annex B (informative) **People's perception about contented life** **15**
Annex C (informative) **Customer satisfaction research on a brand of beer** **20**
Annex D (informative) **Proportions of nonconforming parts of production lines** **26**
Bibliography **31**

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: www.iso.org/iso/foreword.html.

The committee responsible for this document is ISO/TC 69, *Applications of statistical methods*, Subcommittee SC 7, *Applications of statistical and related techniques for the implementation of Six Sigma*.

Introduction

The Six Sigma and international statistical standards communities share a philosophy of continuous improvement and many analytical tools. The Six Sigma community tends to adopt a pragmatic approach driven by time and resource constraints. The statistical standards community arrives at rigorous documents through long-term international consensus. The disparities in time pressures, mathematical rigor, and statistical software usage have inhibited exchanges, synergy, and mutual appreciation between the two groups.

The present document takes one specific statistical tool (Contingency Table Analysis), develops the topic somewhat generically (in the spirit of International Standards), then illustrates it through the use of several detailed and distinct applications. The generic description focuses on the commonalities across studies designed to assess the association of categorical variables.

The Annexes containing illustrations do not only follow the basic framework, but also identify the nuances and peculiarities in the specific applications. Each example will offer at least one “winkle” to the problem, which is generally the case for real Six Sigma and other fields application.

Statistical methods for implementation of Six Sigma — Selected illustrations of contingency table analysis

1 Scope

This document describes the necessary steps for contingency table analysis and the method to analyse the relation between categorical variables (including nominal variables and ordinal variables).

This document provides examples of contingency table analysis. Several illustrations from different fields with different emphasis suggest the procedures of contingency table analysis using different software applications.

In this document, only two-dimensional contingency tables are considered.

2 Normative references

There are no normative references in this document.

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 3534-1 and ISO 3534-2 and the following apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— IEC Electropedia: available at http://www.electropedia.org/

— ISO Online browsing platform: available at http://www.iso.org/obp

3.1
categorical variable
variable with the measurement scale consisting of a set of categories

3.2
nominal data
variable with a nominal scale of measurement

[SOURCE: ISO 3534-2:2006, 1.1.6]

3.3
ordinal data
variable with an ordinal scale of measurement

[SOURCE: ISO 3534-2:2006, 1.1.7]

3.4
contingency table
tabular representation of categorical data, which shows frequencies for particular combinations of values of two or more discrete random variables

Note 1 to entry: A table that cross-classifies two variables is called a "two-way contingency table;" the one that cross-classifies three variables is called a "three-way contingency table." A two-way table with *r* rows and *c* columns is also named "r × c table."

EXAMPLE Let *n* items be classified by categorical variables X and Y with levels X_1, X_2 and Y_1, Y_2, respectively. The number of items with both attribute X_i and Y_j is n_{ij}. Then, a 2 × 2 table is as follows.

Table 1 — 2 × 2 contingency table

Variable X	Variable Y	
	Y_1	Y_2
X_1	n_{11}	n_{12}
X_2	n_{21}	n_{22}

3.5
***p*-value**
probability of observing the observed test statistic value or any other value at least as unfavorable to the null hypothesis

[SOURCE: ISO 3534-1:2006, 1.49]

4 Symbols and abbreviated terms

H_0	null hypothesis
H_a	alternative hypothesis
χ^2	Chi-square statistic
G^2	likelihood-ratio statistic
n	total number of cell count
r × c table	contingency table with r rows and c columns
DF	degree of freedom

5 General description of contingency table analysis

5.1 Overview of the structure of contingency table analysis

This document provides general guidelines on the design, conduct, and analysis of contingency table analysis and illustrates the steps with distinct applications given in Annexes A through D. Each of these examples follows the basic structure given in Table 2.

Table 2 — Basic steps for contingency table analysis

1	State the overall objective
2	List attributes of interest
3	State a null hypothesis
4	Sampling plan
5	Process and analyse data
6	Accept or reject the null hypothesis (Conclusions)

Contingency table analysis is used to assess the association of two or more categorical variables. This document focuses on two-way contingency table analysis, which only considers the relation of two categorical variables. Particular methods for three or more categorical variables analysis are not included in this document. The steps given in Table 1 provide general techniques and procedures for contingency table analysis. Each of the six steps is explained in general in 5.2 to 5.7.

5.2 Overall objectives of contingency table analysis

Contingency table analysis can be employed in Six Sigma[1)] projects in the "Analyse" phase of DMAIC methodologies, and often used in sampling survey, social science and medical research, etc. Apart from the usual statistical methods focusing on continuous variables, contingency table analysis mainly handles the categorical data, including nominal data and ordinal data. In the case that the observed value is the frequency of certain combinations of several objective conditions, but not the continuous value from the equipment, the contingency table analysis is needed.

The primary motivation of this method is to test the association of categorical variables, including the following situations:

a) to assess whether an observed frequency distribution differs from a theoretical distribution;

b) to assess the independence of two categorical variables;

c) to assess the homogeneity of several distributions of same type;

d) to assess the trend association of observations on ordinal variables;

e) to assess extensive association between levels of categorical variables.

5.3 List attributes of interest

This document considers the association of two categorical variables based on the observed frequency of the characteristic corresponding to combinations of different levels of attributes of interest.

If the association between quantitative variable and categorical variable is of interest (e.g. cup size versus surface decoration), it is necessary to divide quantitative data into ordinal classes (e.g. small, medium, large).

5.4 State a null hypothesis

This document is to determine whether row variable and column variable are independent. The null hypothesis for Chi-square test is

H_0: the row variable and column variable are independent;

and the alternative hypothesis is

H_a: the row variable and column variable are not independent.

5.5 Sampling plan

In the sampling plan for contingency table analysis, variables and the levels should be determined first. For two-way contingency tables, there are four possible sampling plans to generate the tables.

a) The total number of cell count n is not fixed.

b) The total number of cell count n is fixed, but none of the total rows or columns are fixed.

c) The total number of cell count n is fixed, and either the row marginal totals or the column marginal totals are fixed;

d) The total number of cell count n is fixed, and both row marginal totals and the column marginal totals are fixed.

---

1) Six Sigma is the trademark of a product supplied by Motorola, Inc. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the product named. Equivalent products may be used if they can be shown to lead to the same results.

The aforementioned four sampling plans correspond to different purposes of categorical data analysis. Case a) is a random sampling, that all frequency numbers are independent. For example, the number of customers entering a supermarket during the day is a random variable. The customers are divided into four classes based on their gender and whether they are shopping or not (male/shopping, male/no shopping, female/shopping, female/no shopping). These four numbers form a contingency table. Case b) is applicable to a sampling survey where the sample size is fixed. Case c) is usually an analysis of a comparative analysis. For example, when conducting a research on the relationship of lung cancer and smoking, a group of patients with lung cancer and a group of healthy people with similar age, gender, and other physical condition are chosen for the research. The total number of people in each group is fixed. Case d) is another test of attribute agreement analysis, usually used to test whether the results from two measurement systems are consistent with each other. For attribute agreement analysis, one can refer to ISO 14468. The calculated statistics of the test of independence for the first three cases are the same.

Randomization is very important when sampling for experiments. The observations in each cell are made on a random sample. When it is inconvenient or difficult to attain adequate samples, one should pay close attention to any confounding factors that may affect the results of the analysis.

Table 3 shows a two-way contingency table with r levels of variable *X* and c levels of *Y*. The observed frequency of each combination of the two variables is n_{ij} (i =1,..., r, j=1,...,c).

Table 3 — Layout of a generic r × c contingency table analysis

Variable *X*	Variable *Y*			
	Y_1	Y_2	...	Y_c
X_1	n_{11}	n_{12}	...	n_{1c}
X_2	n_{21}	n_{22}	...	n_{2c}
...	...	...	...	...
X_r	n_{r1}	n_{r2}	...	n_{rc}

5.6 Process and analyse data

5.6.1 Chi-squared test

Chi-square (χ^2) test is the most fundamental tool for contingency table analysis to test independence of variables. It is commonly used to compare observed data with some expected data according to a specific test purpose.

For a one-dimension contingency table, which has only one categorical variable with two or more levels, Chi-square test, usually called “goodness-of-fit test,” can be used to assess whether the observed data classified by levels follow an theoretical distribution.

For a two-dimensional contingency table, r × c table, Chi-square test can be used to evaluate whether two categorical variables are independent. It can test the homogeneity of distributions with same type, which is also called “homogeneity test.”

Chi-square test is defined to evaluate the distance of the observed data from the expected data. The formula for calculating Chi-square statistic is:

$$\chi^2 = \sum \frac{(o-e)^2}{e} \tag{1}$$

where

o is the observed frequency data;

e is the expected frequency data.

The formula is the sum of the squared difference between observed and expected frequency, divided by the expected frequency in all cells.

For r × c table, there are r levels for row variable X, c levels for column variable Y. With the null hypothesis H_0, X and Y are independent, and the alternative hypothesis H_α, X and Y are not independent, the Chi-square statistic is calculated as follows:

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(n_{ij} - m_{ij})^2}{m_{ij}} \tag{2}$$

where

n_{ij} is the observed frequency in the ith level of row variable X and jth level of column variable Y;

m_{ij} is the expected value of n_{ij} assuming independence.

This statistic takes its minimum value of zero when all $n_{ij} = m_{ij}$. For a fixed sample size, greater differences between n_{ij} and m_{ij} produce larger χ^2 values and stronger evidence against H_0. The χ^2 statistic follows a asymptotic Chi-square distribution for large n, with degree of freedom (DF) = (r-1)(c-1). Reject the null hypothesis if the p-value is less than the pre-specified value, commonly taken at 0,05. The Chi-square approximation improves as m_{ij} increases. Note that when any cell expected value is less than 5, the Chi-square test is not appropriate.

An alternative method for independence test is using likelihood-ratio function through the ratio of two maximum functions. For r × c table, the likelihood functions are based on multinomial distribution, and the likelihood-ratio statistic is

$$G^2 = 2\sum_{i=1}^{r} \sum_{j=1}^{c} n_{ij} \log\left(\frac{n_{ij}}{m_{ij}}\right) \tag{3}$$

This statistic asymptotically follows the χ^2 distribution. The two methods usually have the same properties and provide same conclusions.

If n is small, the distribution of statistics χ^2 and $G^2$are less Chi-squared. Usually, if there exists at least one expected value less than 5 (in some statistical software, the criterion is slightly different), Fisher's exact test is used. In this case, the observed data follow the hypergeometric distribution, which is the exact distribution of the data. The calculation is much more complicated than χ^2 and G^2 where statistical software is often used. There is another way to handle the case when the expected value is too small. One can combine the row or column to the adjacent one to increase the expected values before the independence test. However, this method should be used with caution; combining columns/rows may reduce interpretability and also may create or destroy structure in the table. If there is no clear guideline, the combination method should be avoided.

It should be noted that Chi-square test is basically a test of significance. It can help decide if a relationship exists, but not how strong it is. Chi-square test is not a measure of strength of association.

In this document, Chi-squared test is used to assess three types of comparisons: test of goodness-of-fit, test of independence, and test of homogeneity.

a) Test of goodness-of-fit compares the observed values and theoretical distribution to determine whether an experimenter's prediction fit the data.

b) Test of independence determines whether there is a significant association between two categorical variables.

c) Test of homogeneity compares two sets of categories to determine whether the two groups are distributed differently among the categories.

5.6.2 Linear trend test

When both row and column variables are ordinal data, linear trend test can be used to test whether a trend exists when a variable changes. However, binary variable can also be treated as ordinal data (e.g. no heart disease as "low" risk condition versus having heart disease as "high" risk condition). There are two directions of the trend. Given two ordinal variables *X* and *Y*, as the level of *X* increase, response on *Y* tends to increase to higher levels (positive linear trend), or response on *Y* tends to decrease to lower levels (negative linear trend).

Correlation coefficients are sensitive to linear trends. Two common used correlation coefficients are Pearson's r and Spearman's rho. Pearson's correlation is a measure of linear relationship between two variables. The calculation of Pearson's r is based on the number of observations. Spearman's rho is a nonparametric statistic, using rank orders instead of observations.

Other three correlation coefficients assessing the trend association are Goodman and Kruskal's γ, Kendall's τ (tau-b), and Somers' D. These nonparametric methods depend on the scores of the data in the rows or columns, but not on the quantitative value of each cell.

The values of correlation coefficients range from −1 to 1. The closer the value to 1, the more positive the linear trend is; the closer the value is to −1, the more negative the linear trend is. If the value is 0, there is no relationship. Kendall's τ, like Goodman and Kruskal's γ, tests the significance of association between ordinal variables, and it includes tied pairs (identical pairs) in its calculation, which γ ignores. For Somers' D, D R|C and D C|R measure the strength and direction of the relationship between pairs of variables, with row variable and column variable as the response variables, respectively.

These coefficients just give a result of the possibility of trend association, but the trend association test shows strength of the relationship. For each coefficient, calculate the test statistic z ($z = \gamma$, τ, or d) and its standard error σ_z. The statistic is

$$U = \frac{z}{\sigma_z} \tag{4}$$

with the null hypothesis H_0 variables *X* and *Y* are independent. *U* follows the asymptotic standard normal distribution when H_0 is true. Reject the null hypothesis if the *p*-value is less than the pre-specified value, commonly taken to be 0,05.

Loglinear Model is another useful method for contingency table analysis, which is not included in this document. The models specify how expected cell counts depend on levels of categorical variables and allow for analysis of association and interaction patterns among variables. It often uses for high dimensional tables. Loglinear Model method contains quite a few calculations, which is usually with the aid of computer.

5.6.3 Correspondence analysis

Correspondence analysis (CA) is a statistical visualization method to analyse the association between levels of row and column variables in a two-way contingency table. This technique transforms the rows and columns as points in a two-dimensional space, such that the positions of the row and column point are consistent with their association in the table.

CA changes a data table into two sets of new variables called "factor scores" (obtained as linear combination of row and columns, respectively). The factor scores represent the similarity of the rows and columns structure. Plot the factor scores in a plane that optimally displays the information in the original table. This plot is the standard symmetric representation of CA.

CA is intimately related to the independence Chi-square test. The total variance (often called "inertia") of the factor scores is proportional to independence χ^2 statistic and the factors scores in CA decompose this χ^2 into orthogonal components.

In a symmetric CA plot, the distance between two row (respectively column) points in the coordinate plane is a measure of similarity with regard to the pattern of row (respectively column) relative frequency data. When two row (respectively column) points are close to each other, it shows that there points have similar profiles. However, the positions of row and column points with respect to each other are not directly comparable and should be interpreted with caution. In this case, use asymmetric plot instead. In the asymmetric plot, the distance from a row point to a column point reflects their association directly.

5.7 Conclusions

For the independence test, when *p*-value is less than the prescribed significance level, it shows a significant association between the variables, and the null hypothesis that the two variables are independent should be rejected.

When calculating correlation coefficients for ordinal variables, the positive coefficient z ($0 < z \leq 1$) implies that variables have a positive trend; else, the negative z ($-1 \leq z < 0$) shows a negative trend. From the linear trend test that evaluates the strength of positive trend or negative trend, small *p*-value (less than the significance level) indicates there exists evidence for the positive trend or negative trend.

Correspondence plot shows the association of different levels of rows and columns. In general, the closer the points are, the more similarity between them.

6 Description of Annexes A through D

Four distinct examples of contingency table analysis are illustrated in the annexes, which have been summarized in Table 3 with the different aspects indicated.

Table 3 — Example summaries listed by annex

Annex	Example	Contingency table analysis details
A	Distribution of number of technical issues found after product release to the field	Goodness-of-fit test, JMP
B	What is your perception about contented life?	3 × 5 table, independence test, correspondence test, Minitab
C	Customer satisfaction research on beer	3 × 3 table, independence test, Fisher's exact test, linear trend test, correspondence test, SAS
D	Proportions of nonconforming parts of production lines.	4 × 2 table, independence test, Q-DAS

Annex A
(informative)

Distribution of number of technical issues found after product release to the field[2)]

A.1 General

In the telecommunications industry, new platforms such as a new mobile device platform or a new base transceiver station (BTS) platform follow a very complex and lengthy pre-launch testing process before their “release” in order to optimize their success and their performance once introduced in the field. Many products are likely to be released for a given platform, with each one of these products going through a pre-launch process as well, somewhat simpler and shorter than the pre-launch process of a new platform, but still quite extensive. For a given product, many revisions consisting of small improvements to the product will also be launched according to a specified pre-launch process. This complex multilevel pre-launch process is very well documented and the corresponding products, be they a new platform, a new product within a platform, or a revision of hardware or software within a product, go through considerable testing (verification, validation, regression testing, etc.) before they are released.

The effectiveness and efficiency of the pre-launch process is assessed after release by the number of issues reported by customers to the technical support centre. Once a product is released, it is closely monitored during a period of one year in order to capture and fix as quickly as possible any issues that may have been overlooked during the pre-launch process. Thus, appropriate staffing needs to be allocated for that purpose. “Lessons learned” are developed as to why issues were missed, as well as why issues were introduced in the first place, and actions to prevent them in the future are taken.

At any point in time, there are around 200 products, all within one year from their release date that are being monitored simultaneously by the technical support centre. Every technical issue found on these products is classified as a level A, B, or C, with level A being the highest priority and likely to cause a “stop shipment,” level B being the next priority, and level C being the lowest priority as that category consists of issues not affecting the user directly. The number of technical issues of level A and B is tracked on an hourly basis and its distribution is modelled to provide information on the level of staffing required by the centre to ensure satisfaction of the customers.

A.2 Attributes of interest

The data of interest consists of the number of technical issues of level A and B received and logged in by the technical customer centre on an hourly basis. Only technical issues associated with products launched for less than a year are of interest in this study — it is rare to find such important technical issues visible to the user after the first year’s release. This data is needed to estimate the staffing of the technical centre to ensure speedy and complete resolution of the customer’s issues.

A.3 Hypotheses

Our null hypothesis is that the hourly number of technical issues logged in follows a Poisson distribution with mean and standard deviation to be estimated from the data. A Poisson distribution is a very commonly used distribution to model number of defects that arrive during a specified time interval.

H_0: Number of technical issues A/B received per hour by the technical centre follows a Poisson distribution.

2) This case study was provided by Dr. Michèle Boulanger, JISC-Statistics, Inc.

H_a: Number of technical issues A/B received per hour by the technical centre does not follow a Poisson distribution.

A.4 Sampling plan

The technical customer centre provided the log of all the technical issues that were reported during its 10 h × 6 d weekly operation. For this study, the “sample” consists of all the technical issues of level A/B received over the last two weeks (120 h) and we assume that these two weeks are a “representative” sample of the past weeks and of the future weeks to come.

A.5 Raw data

The data are shown in Table A.1.

Table A.1 — Number of first year post-release technical issues A/B logged in per hour

Number of first year post-launch technical issues A/B per hour during validation week	Frequency (Number of hours)
0	96
1	11
2	6
3	1
4	3
5	0
6	2
7	1
8	0
9	0
10	0

A.6 Analysis

The analysis of the data in Table A.1 was carried out with SAS JMP Pro Version 11.

Summary statistics are provided in Table A.2.

Table A.2 — Summary statistics

Mean	0,475
Std Dev	1,249 958
Std Err Mean	0,114 105
N	120

It is useful to look at the summary statistics, in particular at the mean (0,475) and standard deviation (1,25). A Poisson distribution has the characteristic that its mean and variance are equal. Here, the sample variance is 1,562 4, somewhat larger than the mean, and this raises a question as to whether the Poisson distribution will be a good fit for the data.

Figure A.1 shows the fitted Poisson distribution to the data, with mean setup to be equal to the sample mean. The red line is the fitted distribution.

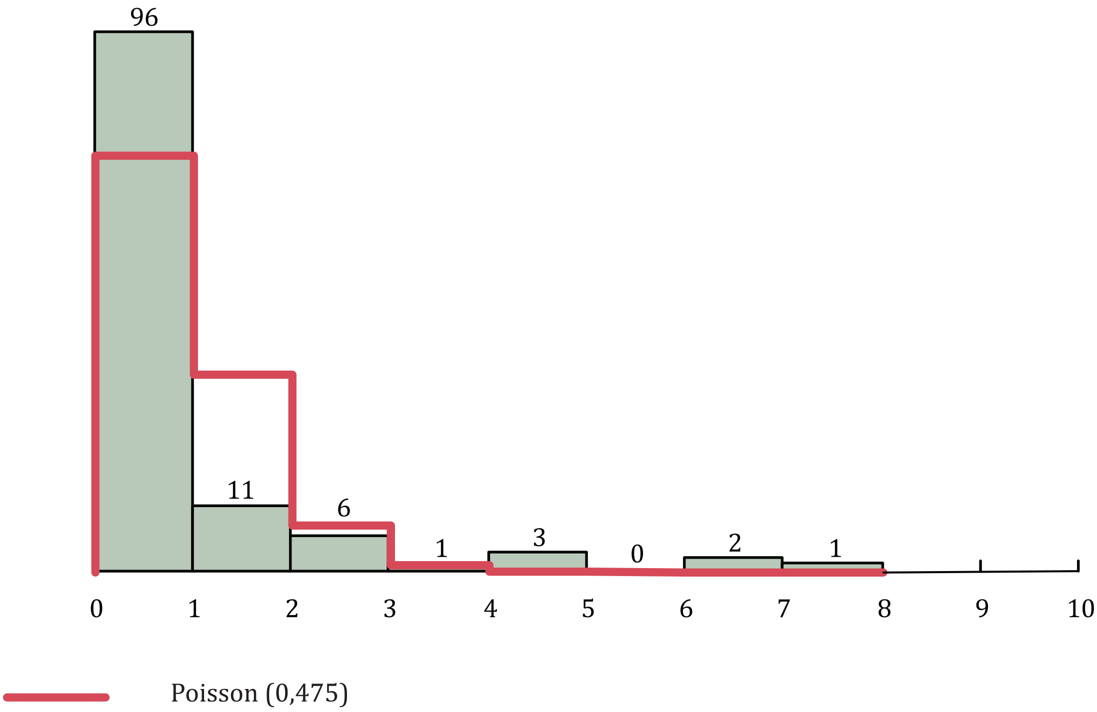


Figure A.1 — Histogram of number of technical issues A/B per hour and fitted Poisson distribution

Table A.3 —Number of technical issues A/B per hour and expected frequencies

Number of first year post-launch technical issues A/B per hour during validation week	Frequency (Number of hours)	Expected frequencies for Poisson
0	96	74,63
1	11	35,45
2	6	8,42
3	1	1,33
4	3	0,16
5	0	0,02
6	2	0,00
7	1	0,00
8	0	0,00
9	0	0,00
10	0	0,00

There are a few cells in Table A.2 with expected frequencies less than 5. To ensure the validity of the Pearson Chi-square test of goodness of fit, regroup the cells with low cell count as shown in Table A.3.

Table A.4 — Consolidated number of technical issues A/B per hour and goodness-of-fit calculations for Poisson distribution

Consolidated number of first year post release technical issues A/B per hour	Frequency	Expected probability with consolidated cells	Expected frequencies for Poisson	Chi-square elements	Chi-square statistic	*p*-value
0	96	0,622	74,63	6,12	—	—
1	11	0,295	35,45	16,86	—	—
2+	13	0,083	9,93	20,70	—	—
—	—	—	—	—	43,69	<0,001

A.7 Conclusions

The null hypothesis is rejected and the Poisson distribution should not be used to make predictions about staffing level or improvement in the launch process.

There are other distributions that can be fitted to these data. In particular, in the case of count type of data and overdispersion (variance larger than the mean, which is the situation in this Appendix), the Gamma Poisson distribution, also known as the "negative binomial distribution," often is a better fit than the Poisson distribution as it arises from a mixture of Poisson distributions, the means of which follow a Gamma distribution. If the number of technical A/B issues from each product follows a Poisson distribution but the means of the Poisson distributions for different products follow a Gamma distribution, then the resulting distribution of technical A/B issues which is being observed will follow a Gamma Poisson distribution.

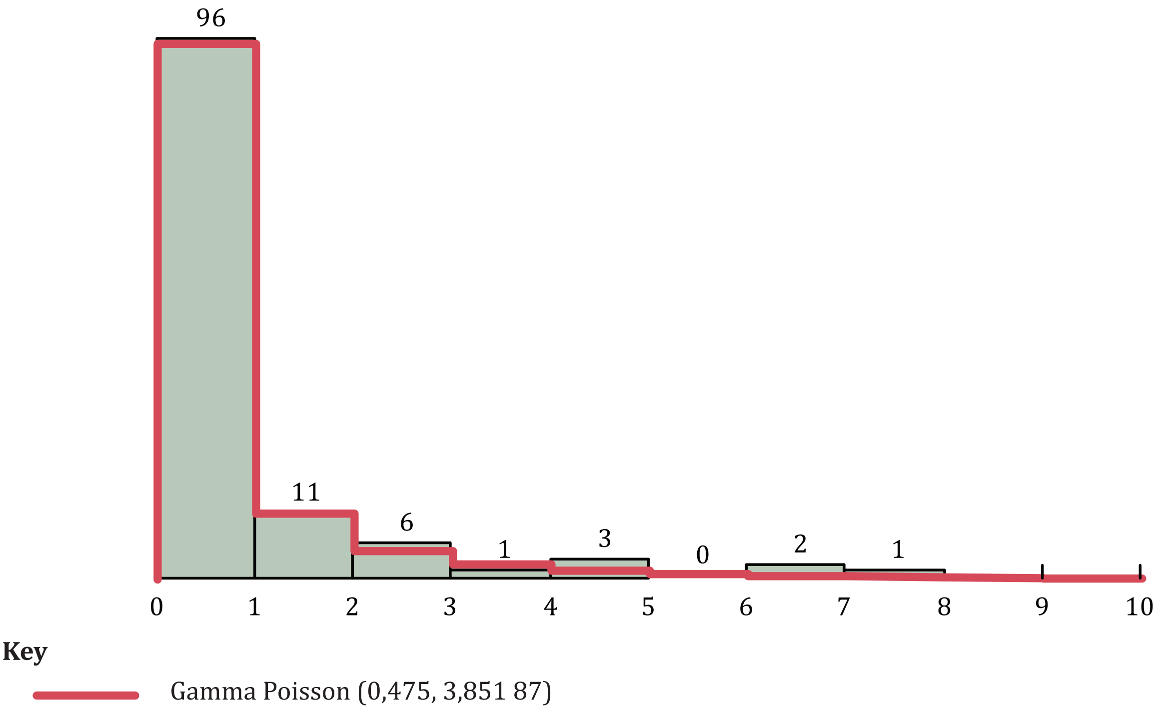


Figure A.2 — Histogram of number of technical issues A/B logged in per hour and fitted Gamma Poisson distribution

There are a few cells in Table A.2 with expected frequency less than 5. To ensure the validity of the Pearson Chi-square test of goodness of fit for the Gamma Poisson distribution, we regrouped the cells with low cell count as shown in Table A.5.

Table A.5 — Consolidated number of technical issues A/B per hour and goodness-of-fit calculations for Gamma Poisson distribution

Consolidated number of first year post-release technical issues A/B per hour	Frequency	Expected probability with consolidated cells	Expected frequencies for Gamma Poisson	Chi-square elements	Chi-square statistic	*p*-value
0	96	0,799	95,86	0,000	—	—
1	11	0,099	11,82	0,057	—	—
2	6	0,043	5,10	0,157	—	—
3+	7	0,060	7,21	0,006	—	—
	—	—	—	—	0,221	0,638

This time, the null hypothesis, H_0: The data is from the Gamma Poisson distribution, is not rejected and thus, it can be used for staffing purposes. As noted previously, the Gamma Poisson distribution is also known as “negative binomial distribution.”

Table A.6 — Parameter estimates

Type	Parameter	Estimate	Lower 95 %	Upper 95 %
Location	λ	0,475	0,287 350 2	0,824 818 2
Overdispersion	σ	3,851 874 4	2,321 399 5	8,015 828 6
2log(Likelihood) = 197,831 173 887 326				

From Table A.6, the data follows the Gamma Poisson distribution with λ = 0,475 and σ = 3,85.

One more verification needs to be done before adopting the Gamma Poisson distribution for our hourly number of technical issues A/B. It relates to the assumption made regarding the “representativeness” of our two weeks’ sample for our data.

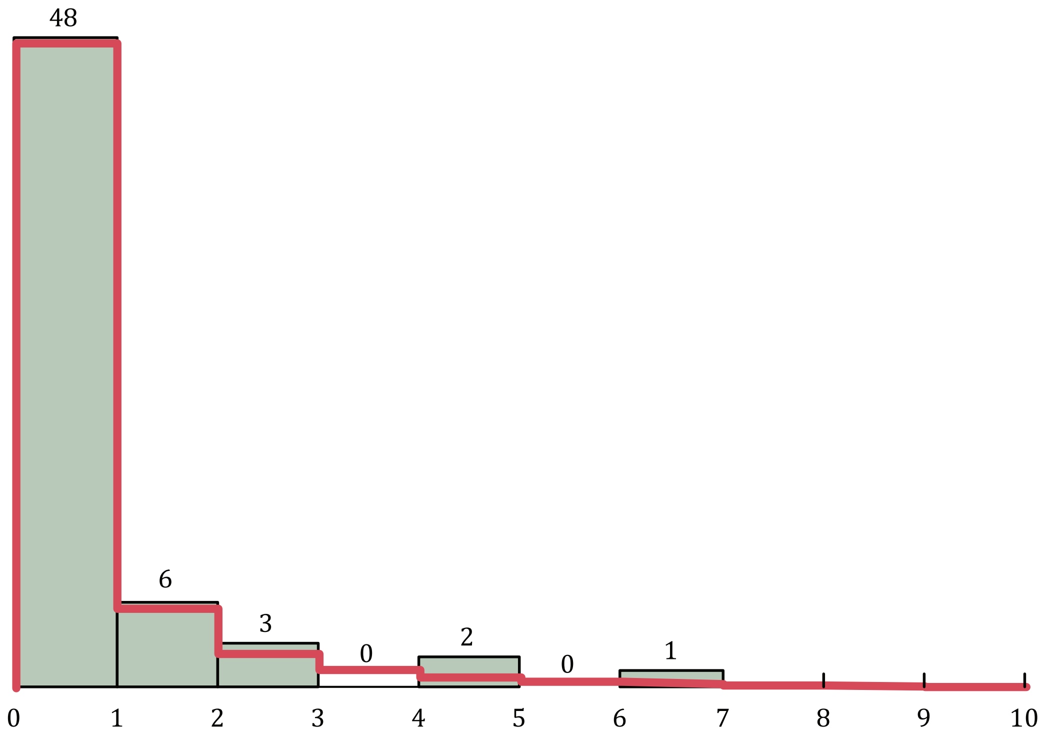
Data were collected for one more week (60 h) to assure that the distribution of technical issues recorded per hour is stable (similar to our two-week analysis period). Data is given in Table A.7.

Table A.7 — Frequencies of technical issues A/B per hour for one additional week of 60 h

Number of first year post launch technical issues A/B per hour during validation week	Frequency (Number of hours)
0	48
1	6
2	3
3	0
4	2
5	0
6	1
7	0
8	0
9	0
10	0

Table A.8 — Summary statistics

Mean	0,433 333 3
Std Dev	1,125 462 9
Std Err Mean	0,145 296 6
N	60



Key

Gamma Poisson (0,475, 3,851 87)

Figure A.3 — Histogram of number of technical issues A/B logged in per hour for one additional week of 60 h and fitted Gamma Poisson distribution

Figure A.3 shows data of the additional week fit Gamma Poisson (0,475, 3,851 87) very well. It is not enough to make the conclusion that the two sets of data follow the same distribution. The test can be carried out by using the likelihood ratio test as follows:

Table A.9 — Parameter estimates (from the additional data)

Type	Parameter	Estimate	Lower 95 %	Upper 95 %
Location	λ	0,433 333 3	0,214 826 2	0,960 899 7
Overdispersion	σ	3,315 128 8	1,751 931 1	9,907 578 1
−2log(Likelihood) = 95,969 897 007 548 6				

Table A.10 — Parameter estimates (from the combined data)

Type	Parameter	Estimate	Lower 95 %	Upper 95 %
Location	λ	0,461 111 1	0,306 933 2	0,713 208 9
Overdispersion	σ	3,671 538 9	2,410 920 8	6,488 796
−2log(Likelihood) = 293,883 609 376 736				

The likelihood ratio statistic is calculated as follows:

G = 293,883 609 376 736−(197,831 173 887 326+95,969 897 007 548 6) = 0,082 538 48

The p-value is evaluated in the Chi-square distribution with 2 degrees of freedom. Large values being critical for the null hypothesis, the p-value is 0,959 570 7 = 0,96.

This likelihood ratio test tests the null hypothesis that the parameters of the Gamma Poisson distributions of the two samples are the same. The Chi-square distribution has 2 degrees of freedom because the number of unknown parameters is reduced by 2 from 4 to 2 unknown parameters.

In conclusion, the estimated Gamma Poisson distribution with fixed parameters 0,475 and 3,851 9 fit the data collected during the verification week well and we can use that distribution for staffing purposes, as well as for assessing the improvement made to the pre-launch process that will be made in the future.

Annex B
(informative)

People's perception about contented life[3)]

B.1 General

An organization named "Manashakti Research Centre," located in Lonawala (about 100 kms from Mumbai, details of the address are 76, Mumbai- Pune Road, Varsoli Lonavla 410401, Maharashtra, India), west part of India, works toward the wellness of the society and has programs for pregnant mothers/fetus, children, and adults of all ages. It is run without donations and government support. Volunteers of the organization give their time, and families who attend the workshops pay nominal fees. There are about 80 families who have dedicated their lives to the activities of the centre. The MRC recently carried out a survey of various age groups to investigate people's perception of contented life.

B.2 Attributes of interest

In this survey, one of the questions asked to participants in age groups 25 to 40, 41 to 60, and 60 above was

What is your perception about contented life?

A) Plentiful money

B) Sterling standard of living

C) Peacefulness in family

D) Prestige

E) Accomplishment of responsibility/spiritual progress

The number of each choice among different age groups is recorded.

B.3 Raw data

Table B.1 lists the raw data.

Table B.1 — Cross-classification of people's perception about contented life

Age group	Code				
	A	B	C	D	E
25 to 40	25	69	538	11	12
41 to 60	17	49	467	7	85
61 and above	6	29	271	5	57

B.4 Contingency table analysis

Contingency table analysis is conducted by Minitab 16[4)].

3) This case study was provided by Dr. Avinash Dharmadhikari, Tata Motors.

4) Minitab is the trade name of a product supplied by Minitab, Inc. This information is given for the convenience of

B.4.1 Independence test

Chi-square test is used to test whether there exists association between age and their perception about contented life.

Null hypothesis H_0: people's age and their perception about contented life are independent.

Alternative hypothesis H_a: people's age and their perception about contented life are not independent.

Statistics in Table B.2 shows the count, expected count, and Chi-square for each cell.

Table B.2 — Tabulated statistic

	A	B	C	D	E	All
25 to 40	25 19,08 1,8385	69 58,43 1,9139	538 507,15 1,8769	11 9,14 0,3779	12 61,21 39,560 2	655 655,00 —
41 to 60	17 18,20 0,0796	49 55,75 0,8171	467 483,92 0,5916	7 8,72 0,3402	85 58,40 12,111 1	625 625,00 —
61 and above	6 10,72 2,077 1	29 32,83 0,445 8	271 284,93 0,681 2	5 5,14 0,003 6	57 34,39 14,868 0	368 368,00 —
All	48 48,00 —	147 147,00 —	1 276 1 276,00 —	23 23,00 —	154 154,00 —	1 648 1 648,00 —
Cell Contents:	Count Expected count Contribution to Chi-square					
Pearson Chi-square = 77,583, DF = 8, *p*-value = 0,000 Likelihood ratio Chi-square = 93,506, DF = 8, *p*-value = 0,000						

Chi-square value is 77,583, and *p*-value is 0,000 is less than the significance level 5 %, which shows that the null hypothesis H_0 should be rejected, i.e. people's age and their perception about contented life are not independent.

B.4.2 Correspondence analysis

Table B.3 gives the inertia and Chi-square decomposition results.

Table B.3 — Chi-squared decomposition

Axis	**Inertia**	**Percent**	**Cumulative Percent**
1	0,046 6	0,989 9	0,989 9
2	0,000 5	0,010 1	1,000 0
Total	0,047 1	—	—

The total inertia is 0,047 08, and 98,99 % of the total inertia can be interpreted by dimension 1.

users of this document and does not constitute an endorsement by ISO of the product named. Equivalent products may be used if they can be shown to lead to the same results.

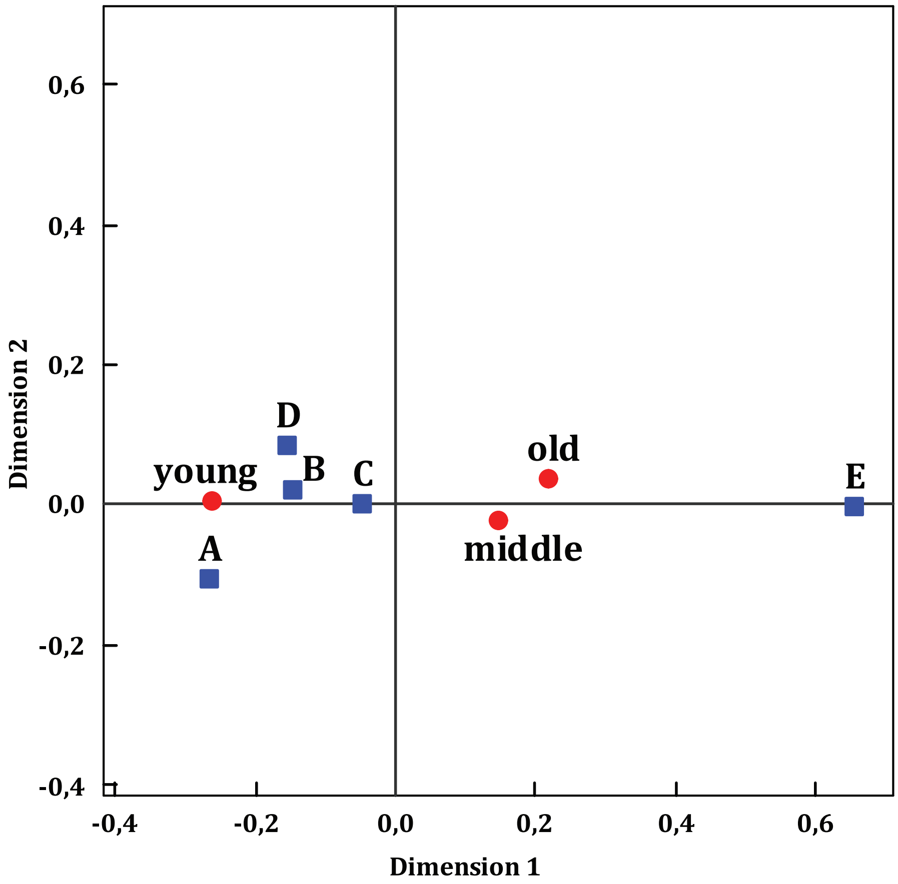
Table B.4 — Row contributions

ID	Name	Dimension 1	Dimension 2
		Coordinate	Coordinate
1	Age 25 to 40	−0,264	0,003
2	Age 41to 60	0,147	−0,024
3	Age 61 and above	0,219	0,034

Table B.5 — Column contributions

ID	Name	Dimension 1	Dimension 2
		Coordinate	Coordinate
1	A	−0,268	−0,108
2	B	−0,146	0,020
3	C	−0,050	0,001
4	D	−0,156	0,084
5	E	0,657	−0,006

Table B.4 and B.5 show the coordinates of row and column levels. Plotting the points in a two-dimensional plane, Figure B.1 is obtained.

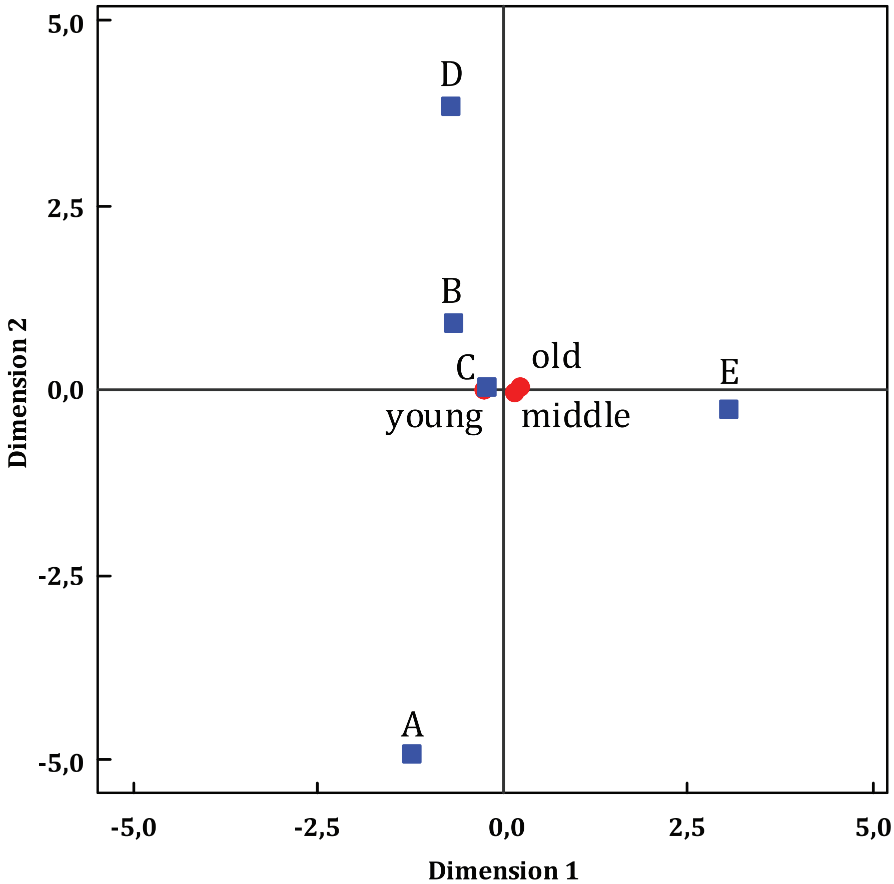


Key

young	age 25 to 40
middle	age 41 to 60
old	age 61 and above
A	plentiful money
B	sterling standard of living
C	peacefulness in family
D	prestige
E	accomplishment of responsibility/spiritual progress

Figure B.1 — Symmetric correspondence analysis plot

The correspondence analysis plot in Figure B.1 shows that old and middle-aged people have similar perception of contented life, which is different from young people. Answer E is isolated from all other answers.



Key

young	age 25 to 40
middle	age 41 to 60
old	age 61 and above
A	plentiful money
B	sterling standard of living
C	peacefulness in family
D	prestige
E	accomplishment of responsibility/spiritual progress

Figure B.2 — Asymmetric correspondence analysis plot

Figure B.2 implies that young people have slightly different perception of contented life from middle-aged and old people. All of them choose relatively more peacefulness in family, sterling standard of living, and prestige. While middle-aged and old would consider accomplishment of responsibility/spiritual progress more compared with young people.

B.5 Conclusions

The Minitab results above indicate that people in different age groups have different perception of contented life. Young and older people have different considerations. Young people usually consider family, sterling standard of living, and prestige important in their life, but for people older than 40, accomplishments of responsibility/spiritual progress is relatively considered more important.

Annex C
(informative)

Customer satisfaction research on a brand of beer[5)]

C.1 General

The degree of customer satisfaction with a particular brand of beer is researched in China. This research is based on a survey on 243 people of age older than 18. All the interviewees answered several questions, including his/her age, degree of satisfaction with the beer, etc. This sampling survey would help the company to collect information of customers' preference with regard to different age. The relationship of customers' age and their satisfaction degree will help to make further targeted strategy to win more market.

C.2 Attributes of interest

To investigate Chinese customer satisfaction with the beer, the questionnaire contains the age of the interviewees and the degree of the satisfaction. For the age variable, it is classified to three categories: young (age 18 to 39), middle (age 40 to 59), and old (age 60 and above). For the levels of satisfaction, it is classified to three levels: low, medium, and high. The response variable is counting how many interviewees in each age category are satisfied with the beer at every level. There are nine possible combinations of the counting. The objective of the study is to explore whether there is an association between customers' age and their satisfaction degree with the beer and further for the correspondence relations.

C.3 Design a sampling plan for study

The customer satisfaction investigation is carried out by telephone survey. The total number of interviewees is not fixed. All the people are selected randomly to answer the questionnaire in a certain period of time.

C.4 Raw data

Table C.1 lists the data collected.

Table C.1 — Cross-classification of customer' satisfaction with the beer

Age	Satisfaction Evaluation		
	low	**medium**	**high**
Young	3	45	86
Middle	1	18	61
Old	3	3	23

5) This case study was provided by Ms. Jian Kang, China National Institution of Standardization.

C.5 Contingency table analysis

Contingency table analysis in SAS[6)] is used to assess the independence and linear trend of the categorical variables, as well as the correspondence analysis producing a visual representation of the relationships between the row and column categories.

C.5.1 Independence test

Chi-square test is adopted to test the independence of the row and column variables.

Null hypothesis H_0: customers' age and their satisfaction degree are independent.

Alternative hypothesis H_a: customers' age and their satisfaction degree are not independent.

Statistics in Table C.2 shows the frequency, expected frequency, and Chi-square for each cell. Old people of low and medium evaluation for the beer have the two biggest cell Chi-square contributions, which implies the satisfaction degree of old people may be different from young and middle-aged people.

Table C.2 — Tabulated statistics

Age	**Evaluation**			**Total**
Frequency **Expected Cell** **Chi-square**	**Low**	**Medium**	**High**	
Young	3 3,860 1 0,191 6	45 36,395 2,034 5	86 93,745 0,639 9	134
Middle	1 2,304 5 0,738 5	18 21,728 0,639 8	61 55,967 0,452 6	80
Old	3 0,835 4 5,608 8	3 7,876 5 3,019 2	23 20,288 0,362 5	29
Total	7	66	170	243

Results of independence test are shown in Table C.3. The Chi-square value assesses whether the two variables (customers' age and their satisfaction degree for the beer) are independent of each other. The *p*-value represents the probability of observing the calculated test statistic could have occurred, assuming that the null hypothesis is true. If the *p*-value is less than the pre-specified significance level (alpha), which is often 0,05, the null hypothesis should be rejected. If one rejects the null hypothesis at the 5 % level, this implies that only 5 % of chance would be the supposed process produce a finding this extreme if the null hypothesis is true.

Table C.3 — Chi-squared test

Statistic	DF	Value	Probability
Chi-square	4	13,6873	0,0084
Likelihood ratio Chi-square	4	12,4658	0,0142
Warning: 33 % of the cells have expected counts less than 5. Chi-square may not be a valid test.			

In Table C.3, the Chi-square value is 13,687 3, and the *p*-value is 0,008 4; the likelihood ratio Chi-square is 12,465 8, and *p*-value is 0,014 2. Both two *p*-values are less than 0,05, which represents the null

6) SAS is the trade name of a product supplied by SAS Inc. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the product named. Equivalent products may be used if they can be shown to lead to the same results.

hypothesis is rejected. But note the warning statement: 33 % of the cells have expected counts less than 5. Chi-square may be not a valid test. So, Fisher's exact test should be used in this case (see Table C.4).

Table C.4 — Fisher's exact test

Table Probability (P)	8.282E-06
Pr ≤ P	0,010 6

In Table C.4, *p*-value is 0,010 6, which is less than 0,05. This shows that the null hypothesis should be rejected, i.e. evidence shows that customer's age and their degree of satisfaction with the beer are not independent.

In Table C.2, the cell Chi-square values of old people are much higher than those of customers in other age categories, which shows that old people's preference may be different from middle-aged and young people. To confirm this result, the test is repeated without old people.

Table C.5 — Tabulated statistics (with young and middle-aged interviewees)

Age	**Evaluation**			
Frequency Expected Cell Chi-square	Low	Medium	High	Total
Young	3 2,504 7 0,098	45 39,449 0,7812	86 92,047 0,397 2	134
Middle	1 1,495 3 0,164 1	18 23,551 1,308 5	61 54,953 0,665 3	80
Total	4	63	147	214

Table C.6 — Chi-squared test (with young and middle-aged interviewees)

Statistic	**DF**	**Value**	**Probability**
Chi-square	2	3,414 4	0,181 4
Likelihood ratio Chi-square	2	3,498 9	0,173 9
Warning: 33 % of the cells have expected counts less than 5. Chi-square may not be a valid test.			

Tables C.5 and C.6 give the statistics and Chi-square test results, but still with a warning that 33 % of the cells have expected counts less than 5, so the Chi-square test may not be valid. Further use Fisher's exact test.

Table C.7 — Fisher's exact test (with young and middle-aged interviewees)

Table Probability (P)	0,009 7
Pr ≤ P	0,158 0

Fisher's exact test result shows the *p*-value is 0,158 0, larger than 0,05, which implies that there is evidence of independence between people of age below 60 and their preference of the beer.

The independence test with all customers and the test without old customers show that old people's satisfaction degree with the beer are significantly different from the customers in other age categories.

C.5.2 Linear trend test

Because the row and column variables are both ordinal data, linear trend test can be used for further analysis. Table C.8 shows several correlation coefficients value for ordinal variables, the asymptotic standard error (ASE) for each statistic, and 95 % confidence limits.

A Spearman correlation 0,126 4 suggests a weak association between customers' age and their satisfaction degree. The statistic Gamma is 0,245 0, Kendall's tau-b is 0,120 9 and Somers' D C|R is 0,105 5 — all of them suggest there exists a weak positive trend. In Tables C.9 and C.10, *p*-values in the test for Kendall's tau-b and Somers' D C|R are all less than 0,05, which indicates a weak positive association, i.e. satisfaction with the beer has a weak increase as age increases.

Table C.8 — Statistics for measures of association

Statistic	Value	ASE	95 % Confidence Limits	
Gamma	0,245 0	0,125 6	−0,001 2	0,491 3
Kendall's tau-b	0,120 9	0,060 9	0,001 6	0,240 3
Stuart's tau-c	0,090 7	0,045 7	0,001 1	0,180 3
Somers' D C|R	0,105 5	0,053 2	0,001 2	0,209 7
Somers' D R|C	0,138 7	0,070 1	0,001 3	0,276 1

Table C.9 — Test for Kendall's tau-b (test of H_0: tau-b = 0)

ASE under H_0	0,060 9
Z	1,984 1
One-sided Pr > Z	0,023 6
Two-sided Pr > |Z|	0,047 2

Table C.10 — Test for Somers' D C|D (test of H_0: Somers' D C|R = 0)

ASE under H_0	0,053 2
Z	1,984 1
One-sided Pr > Z	0,023 6
Two-sided Pr > |Z|	0,047 2

C.5.3 Correspondence analysis

Correspondence analysis provides a multidimensional representation of the association between the row and column categories. Table C.11 shows the row profiles, each cell value is obtained by the frequency of the contingency table divided by the row sum. Singular value indicates the association of the levels of row and columns in a two-dimensional plane.

Table C.11 — Row profiles

	High	Medium	Low
Young	0,641 8	0,022 4	0,335 8
Middle	0,762 5	0,012 5	0,225 0
Old	0,821 4	0,071 4	0,107 1

Chi-square in Table C.12 is the same statistic from Chi-square test; the value is 10,090 8. Inertia measures the strength of association of row and columns. The total inertia is 0,041 7, 81,32 % of which can be interpreted by dimension 1, which implies most of the relations can be interpreted by dimension 1.

Table C.12 — Inertia and Chi-squared decomposition

Singular value	Inertia	Chi-square	Percent	Cumulative percent
0,184 15	0,033 91	8,206 2	81,32	81,32
0,088 25	0,007 79	1,884 6	18,68	100,00
Total	0,041 70	10,090 8	100,00	—
Degrees of freedom = 4				

Table C.13 shows that the first dimension distinguishes mostly between young people and the others, which means young people have different evaluation of the beer compared to the others. From Table C.14, category medium stays on the left side of the origin far away from the other two categories.

Table C.13 — Row coordinates

	Dimension 1	Dimension 2
Young	−0,133 8	0,046 5
Middle	0,070 3	−0,121 0
Old	0,439 6	0,123 0

Table C.14 — Column coordinates

	Dimension 1	Dimension 2
Low	0,496 0	0,499 8
Medium	−0,282 9	0,048 9
High	0,092 3	−0,036 6

Tables C.13 and C.14 with row and column coordinates result in a scatter plot. The correspondence analysis plot in Figure C.2 indicates middle-aged customers have high satisfaction, young people have medium satisfaction, and old people have relatively low satisfaction.

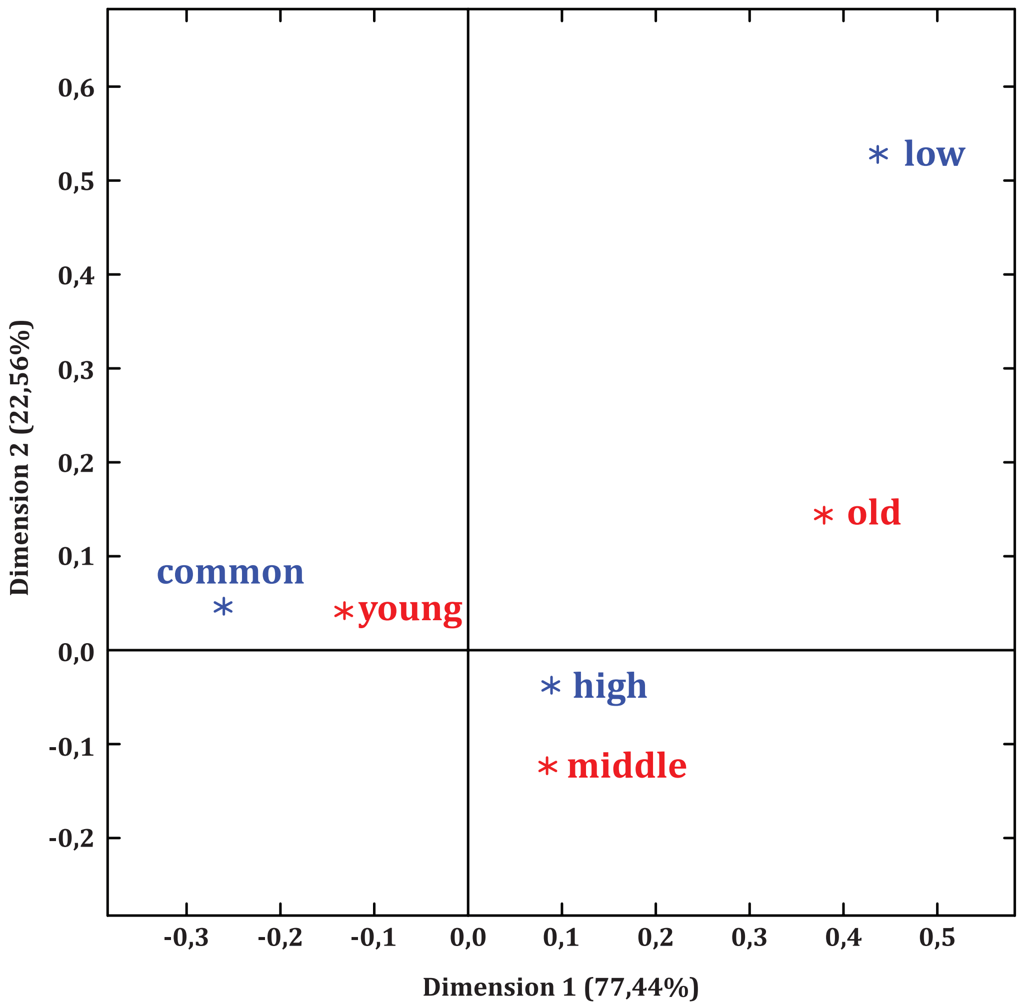


Figure C.1 — Table correspondence analysis plot

C.6 Conclusions

The Chi-square test results show that customers' satisfaction degree has association with the age of customers. There is a weak positive association between the two categories, which is older people incline to have higher evaluation of the beer.

Based on the correspondence analysis plot, young people show medium satisfaction with the beer, middle-aged people show high satisfaction, and old people show relatively low satisfaction.

There is a clearly different preference for the beer among different age customers. Some reasons may lead to above conclusions. Young people easily accept new tastes and brands, and their brand loyalty is not that strong compared with others. While old customers may not usually drink beer in China, they show relatively low satisfaction with the beer. Middle-aged people are the major consumption group. As a conclusion, the difference of customers' satisfaction at some level is caused by the consumption habit of different age groups. The beer company may need to exert more effort to explore the market of young consumers.

Annex D
(informative)

Proportions of nonconforming parts of production lines[7)]

D.1 State the overall objectives

In a factory, electronic control units are produced. Four production lines work in parallel producing these units. The proportion of defective parts for the control units, p, is found to be at $p \approx 5$ %. That value is unacceptably high for this kind of units and needs to be decreased within the framework of a Six Sigma project.

One of the questions to be answered during the "Analyse" phase is whether the proportion of defective units differs between the lines. If this was the case, further analysis concerning differences between worse and better lines can lead to causes for the observed high number of defective parts.

It is assumed that differences in p exist among the lines. In order to confirm that, a χ^2 test for homogeneity is applied. The performance concerning quality for each line is expressed as p_1, p_2, p_3, and p_4 for the proportion of defective parts in lines 1, 2, 3, and 4.

D.2 List attributes of interest

In terms of attributes for a contingency table, two categorical variables are considered. Y variable stands for quality with its two possible levels: defective and not defective. Both levels are mutually exclusive. The X variable describes the production line with levels 1, 2, 3, and 4. For all combinations of X and Y levels, the frequencies of occurrence can be found by recording the number of defective and not defective control units for each production line over a certain period of time. If the number of produced units in each line is designated as n_i with i=1..4, the number of not defective units in line i can be derived from (n_i-y_i). Here, y_i stands for the number of defective units in line i. The parameters p_1, p_2, p_3 and p_4 are estimated by dividing the frequencies of the level 'defective' of the Y variable y_i by the number of units produced n_i for each level i of variable X (the lines).

D.3 State a null hypothesis

The proportion of defective parts for each production line p_1, p_2, p_3, and p_4 are considered to be distribution parameters of four binomial distributions (together with the respective parameters of the sample size n_i i = 1, 2, 3, 4) of four different populations.

The hypotheses are derived from the aim to determine whether or not the parameters p_1, p_2, p_3, and p_4 differ significantly.

Two complementary hypotheses are formulated:

Null hypothesis, H_0: $p_1 = p_2 = p_3 = p_4$

The production lines have all the same proportion of defective units.

Alternative hypothesis, H_a: $p_i \neq p_j$ i,j=1, 2, 3, 4 with $i \neq j$

At least one production line differs in the proportion of defective units from another line.

7) This case study is provided by Mr. René Pleul, TEQ Training & Consulting GmbH.

If H_0 was true, the number of observations in the cells of the contingency table would be homogeneously distributed over the lines following a common proportion $p = p_1 = p_2 = p_3 = p_4$ and the sample size n_i for each line. In other words, no relationship exists between the variable Y (quality as percentage of defective units) and the variable X (the different lines).

D.4 Design a sampling plan for study

During the "Measure" phase of the Six Sigma project, it has been proven (using a control chart) that the proportion of defective control units does not significantly vary over time. Because of that, it is assumed that, in terms of quality, the processes on each production line are stable as well.

Thus, for this study, the number of defective control units y_i is recorded for each production line over a defined period of time. Due to changing workloads on the lines, the total number of produced units is counted as well and gives the marginal totals n_i, i=1..4. The number of not defective units is to be calculated following (n_i - y_i) as mentioned earlier.

Table D.1 shows the contingency table and the acquired data for the analysis.

D.5 Raw data

The bold printed data in Table D.1 have been recorded. The numbers of not defective units has been calculated. For both variables X and Y, a 4 × 2 contingency table was created and the data was filled in.

Table D.1 — Numbers of defective and good units of four production lines

X variable "Production lines"	*Y* variable "Quality"			Estimate of
	Defective units (y_i)	Not defective units (n_i - y_i)	Number of produced units	p_i (i =1, 2, 3, 4)
Line 1	10	117	127	0,787
Line 2	12	144	156	0,769
Line 3	7	133	140	0,050
Line 4	4	253	257	0,016

D.6 Analysis

The comparison of proportions of nonconforming parts requires the discrete binomial distribution (defective/good) as displayed by the software (destra®[8)]). In the present example, four production lines are compared. Thus, the χ^2 (homogeneity) test for more than two populations is to be chosen (see Figure D.1).

8) destra® is the trademark of a product supplied by Destra Software Ltd. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the product named. Equivalent products may be used if they can be shown to lead to the same results.

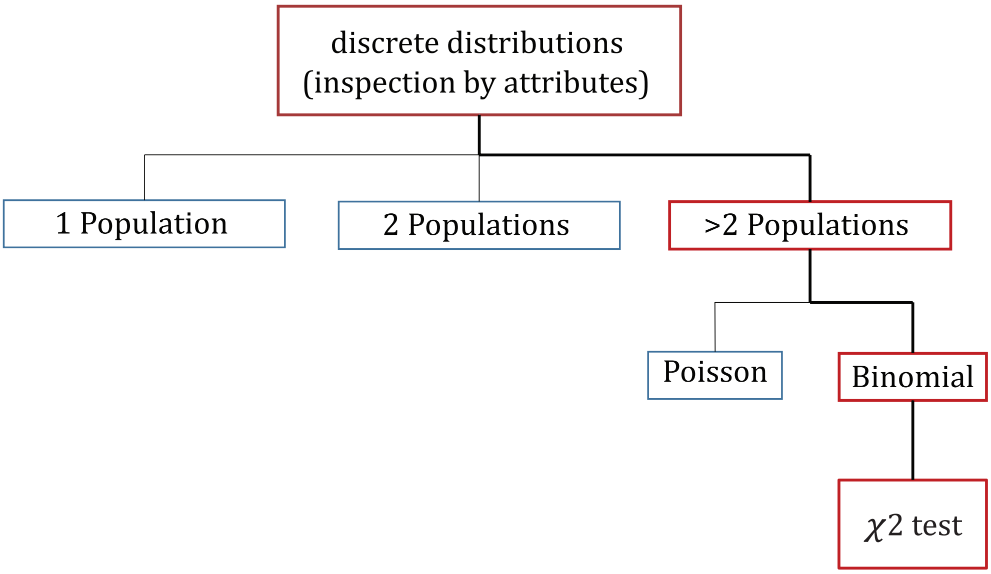


Figure D.1 — Guided selection of χ^2 test in destra®

The result of the χ^2 test in destra® is as follows.

Table D.2 — χ^2 test in destra®

Comparison of expected values of binomial distributions						
p test(BD) k > 1						
H_0	The expected values of the populations are equal					
H_a	The expected values of the populations are NOT equal (at least for one pair)					
Test level	Critical values		Test statistics			
	lower	upper				
α = 5 %	—	7,81	$\sum_l \frac{(x_l - n_l\bar{p})^2}{n_l\bar{p}} + \sum_l \frac{[n_l - x_l - n_l(1-\bar{p})]^2}{n_l(1-\bar{p})} = 11,2891$			
α = 1 %	—	11,34				
α = 0,1 %	—	16,27				
Test results						
Null hypothesis rejected at level $\alpha \leq 5$ %						
Pop.	Active	Description	*n*	*x*	Test factor (*x*)	Test factor (*n-x*)
1	X	production line 1	127	10	2,388 48	0,121 82
2	X	production line 2	156	12	2,591 57	0,132 18
3	X	production line 3	140	7	0,006 238 9	0,000 318 21
4	X	production line 4	257	4	5,754 93	0,293 53

The bottom table in Table D.2 shows the data for the production line in four rows. In column “*x*,” the numbers of defective units y_i are given. In column “*n*,” the total numbers of produced units n_i are entered according to the acquired data. The columns “Test factor (*x*)” and “Test factor (*n-x*)” show the contributions to the test statistic χ^2. This refers to the formula shown in the field above. Following Formula (2), with c = 2 and r = 4, this formula comprises c = 2 summands each for one column in the contingency table [“Test factor (*x*)” for the number of defective units and “Test factor (*n-x*)” for the number of not defective]. Index *l* in both terms concerns *l*=1..*r* different populations and refers to the production line *i* as introduced above.

The expected values for the number of defective units are calculated by $n_i\bar{p}$ for each production line. For the not defective units, it is $n_i(1-\bar{p})$. Here, $\bar{p}$ is the estimation for the common proportion p under

H_0. It is to be calculated by $\bar{p} = \sum y_i \,/\, \sum n_i$. For the given data, $\bar{p}$ =0,048 5. None of the expected values in the example falls below 5. Thus, the χ^2 approximation of the test statistic is accepted. The result for the test statistic [as the sum over the values in the columns "Test factor (x)" and "Test factor (n-x)"] is χ^2 = 11,289.

Instead of using a *p*-value, the final test result is to be concluded by comparing the calculated χ^2-value of χ^2 = 11,289 with the critical χ^2-value derived from the χ^2-distribution for a defined Type I error probability α (and DF, the available degrees of freedom, DF = 3 in the example). For all values of the χ^2-test statistic higher than the critical percentiles $\chi^2_{1\text{-}\alpha,\text{DF}}$, H_0 is to be rejected (this would occur in α% of the cases if H_0 was true). The left upper part in Table D.2 shows the critical values for α = 5 %, 1 % and 0,1 %. These values are chosen for three degrees of significance in rejecting H_0. The comparison result and the conclusion for these three α-values are displayed in Table D.3.

Table D.3 — Comparison of critical χ^2 percentiles and χ^2 test statistic

Test level	Comparison of χ^2 value and critical value	Conclusion	Degree of significance / α-category	
α = 5 %	11,289 > 7,81	H_0 is rejected.	significant *	1 % < α ≤ 5 %
α = 1 %	11,289 < 11,34	H_0 is not rejected.	very significant **	0,1 % < α ≤ 1 %
α = 0,1 %	11,289 < 16,27	H_0 is not rejected.	high significant ***	α ≤ 0,1 %

The *p*-value corresponding to this test is 0,010 27, which is close to the 1 % significance level. Further, the test factors in Table D.1 (i.e. 2,388 48, 2,559 157, 0,006 238 9, and 5,754 93) sum to the χ^2-test statistic value of 11,289 1.

D.7 Conclusion

From Table D.3, as a result of the rejection of H_0 at the 5 % level, the difference in the observed proportions of the four production lines cannot be explained by random. Differences exist in the proportion of defective units between the production lines. Thus, at least two lines can be compared in a further analysis in order to identify causes for producing defective units. The next question would be: which lines should be compared?

It is shown in Table D.2 that the large χ^2 value (5,75; 0,29) for production line 4 stands out and seems to be different from the other production lines.

In order to test this assumption, it needs to be found out whether the proportion of defective units between line 1 to 3 differs without taking line 4 into account. The test is repeated without line 4.

The result is shown in Table D.4.

Table D.4 — χ^2 test in destra® without line 4

Comparison of expected values of binomial distributions			
p test(BD) k > 1			
H_0	The expected values of the populations are equal.		
H_a	The expected values of the populations are NOT equal (at least for one pair).		
Test level	Critical values		Test statistics
	lower	upper	
α = 5 %	—	5,99	$\sum_I \frac{(x_I - n_I \bar{p})^2}{n_I \bar{p}} + \sum_I \frac{[n_I - x_I - n_I(1-\bar{p})]^2}{n_I(1-\bar{p})} = 1,132\ 19$
α = 1 %	—	9,21	
α = 0,1 %	—	13,82	
Test results Null hypothesis rejected.			

Pop.	Active	Description	*n*	*x*	Test factor (*x*)	Test factor (*n*-*x*)
1	X	production line 1	127	10	0,192 06	0,014 136
2	X	production line 2	156	12	0,159 23	0,011 720
3	X	production line 3	140	7	0,703 28	0,051 764
4	—	production line 4	257	4	—	—

From Table D.4, it can be found that the null hypothesis is rejected at all test levels. This indicates that the proportions of defective units for the first three production lines are not different. Since the quality level of line 4 is relatively better than that of the other three lines, it can be concluded that production line 4 produces fewer defective units than the other lines. In the “Analyse” phase of the Six Sigma project, it is desired to bring production lines 1 to 3 to the level of performance of line 4.

Bibliography

[1] ISO 12888, *Selected illustrations of gauge repeatability and reproducibility studies*.

[2] ISO 3534-1:2006, *Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability*

[3] ISO 3534-2:2006, *Statistics — Vocabulary and symbols — Part 2: Applied statistics*

[4] ISO/TR 14468:2010, *Selected illustrations of attribute agreement analysis*

[5] Agresti A. An Introduction to Categorical Data Analysis. John Wiley & Sons, Inc, Second Edition, 2007

[6] MINITAB. INC. *Meet Minitab 16*. Available at: http://www.minitab.com/uploadedFiles/Shared_Resources/Documents/MeetMinitab/EN16_MeetMinitab.pdf

ICS 03.120.30

Price based on 31 pages

This page deliberately left blank

British Standards Institution (BSI)

BSI is the national body responsible for preparing British Standards and other standards-related publications, information and services.

BSI is incorporated by Royal Charter. British Standards and other standardization products are published by BSI Standards Limited.

About us

We bring together business, industry, government, consumers, innovators and others to shape their combined experience and expertise into standards -based solutions.

The knowledge embodied in our standards has been carefully assembled in a dependable format and refined through our open consultation process. Organizations of all sizes and across all sectors choose standards to help them achieve their goals.

Information on standards

We can provide you with the knowledge that your organization needs to succeed. Find out more about British Standards by visiting our website at bsigroup.com/standards or contacting our Customer Services team or Knowledge Centre.

Buying standards

You can buy and download PDF versions of BSI publications, including British and adopted European and international standards, through our website at bsigroup.com/shop, where hard copies can also be purchased.

If you need international and foreign standards from other Standards Development Organizations, hard copies can be ordered from our Customer Services team.

Copyright in BSI publications

Storing and using standards

Standards purchased in soft copy format:

- A British Standard purchased in soft copy format is licensed to a sole named user for personal or internal company use only.
- The standard may be stored on more than 1 device provided that it is accessible by the sole named user only and that only 1 copy is accessed at any one time.
- A single paper copy may be printed for personal or internal company use only.

Standards purchased in hard copy format:

- A British Standard purchased in hard copy format is for personal or internal company use only.
- It may not be further reproduced – in any format – to create an additional copy. This includes scanning of the document.

If you need more than 1 copy of the document, or if you wish to share the document on an internal network, you can save money by choosing a subscription product (see 'Subscriptions').

Reproducing extracts

For permission to reproduce content from BSI publications contact the BSI Copyright & Licensing team.

Subscriptions

Our range of subscription services are designed to make using standards easier for you. For further information on our subscription products go to bsigroup.com/subscriptions.

With **British Standards Online (BSOL)** you'll have instant access to over 55,000 British and adopted European and international standards from your desktop. It's available 24/7 and is refreshed daily so you'll always be up to date.

You can keep in touch with standards developments and receive substantial discounts on the purchase price of standards, both in single copy and subscription format, by becoming a **BSI Subscribing Member**.

PLUS is an updating service exclusive to BSI Subscribing Members. You will automatically receive the latest hard copy of your standards when they're revised or replaced.

To find out more about becoming a BSI Subscribing Member and the benefits of membership, please visit bsigroup.com/shop.

With a **Multi-User Network Licence (MUNL)** you are able to host standards publications on your intranet. Licences can cover as few or as many users as you wish. With updates supplied as soon as they're available, you can be sure your documentation is current. For further information, email subscriptions@bsigroup.com.

Revisions

Our British Standards and other publications are updated by amendment or revision.

We continually improve the quality of our products and services to benefit your business. If you find an inaccuracy or ambiguity within a British Standard or other BSI publication please inform the Knowledge Centre.

Useful Contacts

Customer Services
Tel: +44 345 086 9001
Email (orders): orders@bsigroup.com
Email (enquiries): cservices@bsigroup.com

Subscriptions
Tel: +44 345 086 9001
Email: subscriptions@bsigroup.com

Knowledge Centre
Tel: +44 20 8996 7004
Email: knowledgecentre@bsigroup.com

Copyright & Licensing
Tel: +44 20 8996 7070
Email: copyright@bsigroup.com

BSI Group Headquarters

389 Chiswick High Road London W4 4AL UK