TECHNICAL REPORT

ISO/TR 20693

First edition
2019-04

Statistical methods for implementation of Six Sigma — Selected illustrations of distribution identification studies

Méthodes statistiques pour la mise en œuvre du Six Sigma - Exemples choisis d'études d'identification de la distribution

Reference number
ISO/TR 20693:2019(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

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 of the voluntary nature of standards, 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 www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 69, *Applications of statistical methods*, Subcommittee SC 7, *Applications of statistical and related techniques for the implementation of Six Sigma*.

Any feedback or questions on this document should be directed to the user’s national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

Many statistical techniques assume that the data to be analysed come from a given distribution (or population). Such assumptions are crucial to the effectiveness of subsequent statistical inference methods. In the Six Sigma community, when using such statistical methods, one needs to consider whether this assumption is reasonable. More generally, sometimes it is interesting and necessary to find the distribution which generated the data set (or sample) at hand. Identification of the distribution may provide some ways to answer this question. It consists of finding a distribution (or a family of distributions) which provides a good representation of a sample.

The distribution identification within Six Sigma projects should ideally be performed before the end of the Measure phase and can continue throughout the other phases of the DMAIC. From a Six Sigma perspective, the distribution identification can have multiple purposes based on the considered phase. It is used, for example, to characterise a baseline of the process performance, during the Measure or Analyse phase, to characterise the new process during the Improve phase, and to continuously monitor the process performance during the Control phase to ensure that the change is sustained. From a statistical perspective, distribution identification may be helpful to find appropriate statistical techniques for the related data, since many parametric statistical inference methods need certain distributional assumptions.

In general, distribution identification methods may be used as a tool to:

a) verify that a distribution used historically is still valid for the current data;

b) choose the appropriate distribution.

The choice of appropriate distribution should be guided by the knowledge of physical phenomena or the business process. It is recommended to start from a tentative theory to avoid just curve fitting.

In practice, there is always certain context or business background which can be used in determining the distribution. For example, under some circumstance, one can expect the measurement error is normally distributed. In reliability fields, the life distributions for certain products are exponential, lognormal, Weibull, or extreme distributions and so on. However, when such knowledge is not available, the possible underlying distribution for the data should also be identified if one wants to use parametric statistical methods. In this case, exploratory data analysis methods should be used to gain a better understanding. Through graphical visualisation methods, one could form a hypothesis on the possible distributions, stratification of the data or other aspects. Once the hypothesis is formed, hypothesis testing, including goodness of fit testing, can be applied to check one's guess. Finally, a suitable distribution may be found for the data.

In some commercial software packages including MINITAB[1)], SAS-JMP[1)] and Q-DAS[1)], although there are buttons for distribution identification, one should take knowledge of context and process related to data into consideration instead of simply relying on the software packages. Otherwise, misleading results can be given.

1) MINITAB is the trade name of a product supplied by Minitab Inc. JMP is the trade name of a product supplied by SAS Institute Inc. Q-DAS is the trade name of a product supplied by Q-DAS GmbH. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of these products.

Statistical methods for implementation of Six Sigma — Selected illustrations of distribution identification studies

1 Scope

This document provides guidelines for the identification of distributions related to the implementation of Six Sigma. Examples are given to illustrate the related graphical and numerical procedures.

It only considers one dimensional distribution with one mode. The underlying distribution is either continuous or discrete.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements 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.

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

3 Terms and definitions

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

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

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

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

3.1
population
totality of items under consideration

[SOURCE: ISO 3534-1:2006, 1.1, modified - Notes 1, 2, and 3 deleted.]

3.2
sample
subset of a *population* (3.1) made up of one or more sampling units

[SOURCE: ISO 3534-1:2006, 1.3, modified - Notes 1and 2 deleted.]

3.3
observed value
obtained value of a property associated with one member of a *sample* (3.2)

[SOURCE: ISO 3534-1:2006, 1.4, modified - Notes 1 and 2 deleted.]

3.4
family of distributions
distribution family
set of probability distributions

[SOURCE: ISO 3534-1:2006, 2.8, modified - Synonym "distribution family" added; Notes 1 and 2 deleted.]