Computers & Industrial Engineering 72 (2014) 239–246

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

Computer-aided variables sampling inspection plans for compositional proportions and measurement error adjustment ☆

Haizhen Wu *, Kondaswamy Govindaraju

Institute of Fundamental Sciences, Massey University, Private Bag 11222, Palmerston North 4442, New Zealand

ARTICLE INFO

Article history:
Received 1 August 2013
Received in revised form 13 February 2014
Accepted 28 March 2014
Available online 13 April 2014

Keywords:
Compositional proportion
Measurement error adjustment
Non-normality
Weighted kernel deconvolution
Variables acceptance sampling

ABSTRACT

Many quality characteristics encountered in the food industry are compositional proportions, e.g. protein percentage in milk powder. The distributions of these quality characteristics are intrinsically not normal as well as cannot be well-approximated by it. Moreover the shape of underlying distribution of the quality characteristic in each lot is likely to change in a short-run production process due to process adjustment actions and heterogeneity in raw materials. As a result, the standard variables sampling plans based on the normal distribution are not appropriate. The impact of measurement error, which is often strongly present in analytical testing, has not yet been well-addressed in the non-normal variables sampling inspection procedures. This paper proposes a computer-aided procedure for the identification of the underlying distribution, adjustment for the measurement error, and then the design of the sampling inspection plan. The weighted kernel deconvolution approach is employed for measurement error adjustment and a new procedure for designing a variables plan allowing for uncertainty in the shape parameters of the underlying distribution is developed.

1. Introduction

Compositional proportions such as protein percentage in milk powders are often key quality characteristics in the food industry. The distribution of these quality characteristics are naturally bounded in the unit interval and often effectively supported by a narrower interval. Non-normal distributional features such as skewness and excess kurtosis are also observed in practice. Hence the underlying distribution for compositional characteristics is intrinsically not normal. The normal approximation is also poor. As a result, the normal distribution based variables acceptance sampling procedures given in ISO/FDIS 3951-1:2004 E (2004) are not adequate for compositional proportions.

The effect of non-normality on the performance measures of variables sampling plans such as OC (operating characteristic) curves and (producer's and consumer's) risks is discussed in Das and Mitra (1964), Singh (1966) and more recently in Aslam and Jun (2010). However, the problem of designing non-normal variables acceptance sampling plan, particularly for compositional proportions supported on a bounded interval, is still far from being resolved.

The two popular approaches to tackle non-normality are (i) applying a normalizing-transformation to transform the observed data to normal (as suggested in Section 9.5.8, Stephens (2001) etc.) and (ii) developing sampling plans assuming a known non-normal distribution e.g. Tsai and Wu (2006), Aslam and Jun (2010) etc. The former is based on an implicit assumption that the transformed data follow a normal distribution, which is often invalid. For example, the beta-distributed data can never be transformed to normal by Box-Cox transformation (Box & Cox, 1964), even though the transformed data may sometimes pass the normality tests. For the latter approach, only a few types of distributions, which do not include bounded distributions, provide an analytical expression for OC curves of sampling plans. Moreover, the shape parameters for the distributions are usually assumed known, as shown in other recent papers dealing with the design of non-normal acceptance sampling plans (see Aslam & Jun, 2009, 2013; Aslam, Azam, & Jun, 2013; Aslam, Lio, & Jun, 2013; Aslam, Balamurali, Jun, & Ahmad, 2013; Tsai, Lu, & Wu, 2008, etc.), which is often not in line with the reality.

In this paper, we establish a general framework to handle the performance measurement and design of the variables acceptance sampling plans for quality characteristic allowing for (non-normal) continuous parametric distributions, including but not limited to compositional proportions with bounded distributions. The Monte-Carlo simulated OC curves are employed to overcome the obstacle due to unavailability of their analytical counterpart. A

☆ This manuscript was processed by Area Editor H. Brian Hwarng.

* Corresponding author. Tel.: +64 21 251 5897; fax: +64 6 350 5682.
E-mail addresses: h.wu2@massey.ac.nz (H. Wu), k.govindaraju@massey.ac.nz (K. Govindaraju).

http://dx.doi.org/10.1016/j.cie.2014.03.022
0360-8352/

computer-aided system is developed due to lack of analytical solutions and the concept of estimated OC curves is introduced to cope with the situation of unknown shape parameters.

Another focus of this paper is the adjustment for measurement errors. Although the standard acceptance sampling plans are usually designed without consideration of the measurement errors, they are unavoidably present in practice and may have a significant impact on the performance of the sampling plans. This issue is discussed by many researchers such as Owen and Chou (1983), and a straightforward approach is to adjust the empirical distribution of the observed data such that the acceptance sampling procedures can be conducted based on an error-adjusted empirical distribution. This can be achieved by a variance separation approach (Hahn, 1982), when both quality characteristic and measurement error are normally distributed. However, when the underlying quality characteristic is not normally distributed, a feasible solution has not been established to the best of our knowledge.

Under the commonly used and simplest additive measurement error model, the observed data Y is represented as the sum of true quality characteristic X (with density f) and the measurement error E (with density π). Under the assumption that X and E are independent and π is known, the density g of Y is simply the convolution of f and π. However, for non-normally distributed X, it is rare that analytical form of g can be obtained. Therefore, fitting observed data to the convoluted distribution g to gain the knowledge of f by analytical deconvolution is not feasible in most cases. Hence we may attempt to fit the empirical data to the distribution of X directly but this empirical knowledge about the distribution of X is not directly observable. Inspired by the recent progresses in nonparametric and semi-parametric density deconvolution using weighted kernel estimators (Hazelton & Turlach, 2008, 2010), we propose to use the observed data $\{Y_i\}_{\{1\leqslant i\leqslant n\}}$ in conjunction with an appropriate set of estimated weights $\{\hat{w}_i\}_{\{1\leqslant i\leqslant n\}}$ to describe the empirical knowledge of the error-adjusted (which mimics the true) distribution, such that the distribution of X can be estimated. The acceptance sampling procedures based on this error-adjusted distribution can therefore (at least partially) eliminate the impact of the measurement errors.

The layout of this paper is as follows. Section 2 reviews the estimation of the underlying distribution in a general framework and shows the extra difficulty brought by extending underlying distribution from normal to more general distributions. In Section 3, an algorithm for designing variables sampling inspection plans for non-normal distribution with unknown shape parameter is proposed. The identification of underlying distribution of the compositional proportions is discussed in Section 4. The proposed measurement error adjustment method is presented in Section 5. The final Section 6 contains the concluding remarks.

2. Test statistic and OC curve for unknown underlying distribution

Before trying to establish a general framework of designing variables acceptance sampling plans for (non-normal) continuous parametric underlying distributions, it is necessary to first review the relationships of the test statistic and the OC curve with the underlying distribution. It is also necessary to figure out what new problems arise when the underlying distribution is extended from normal to general non-normal distributions. It is worth pointing out that the "general" distributions we refer to are continuous parametric distributions (with location and/or scale parameters) and this setting is sufficient for most cases in practice.

2.1. Measure of quality

Consider the distribution of a quality characteristic X being $F_X(x;\theta_0)$ with parameters vector θ_0, the lot quality measure is often the proportion nonconforming (or the proportion of defectives) p, which is the probability of X falling into a region of nonconformance R^{NC}. That is, p is given by

$$p = \Pr\left(X \in R^{NC}\right) = F_X\left(R^{NC};\theta_0\right). \tag{1}$$

When a single specification limit (say L, the lower specification) is specified, or equivalently the nonconformance region being $(-\infty, L]$, p becomes a monotone function of the specification limit, i.e.

$$p = F_X(L;\theta_0). \tag{2}$$

2.2. Test statistic

The true proportion nonconforming p is never known because error-free 100% inspection is seldom possible. Therefore, the best knowledge we can get about the lot quality p is an appropriate estimate $\hat{p}$, and this becomes a natural test statistic to judge the quality of the lot. In the context of acceptance sampling, the test statistic leads to acceptance (or rejection) decisions. Consider the decision criteria involving an acceptance constant p^* (also called maximum allowable proportion nonconforming), such that $\hat{p} \leqslant p^*$ ($\hat{p} \geqslant p^*$) becomes the acceptance (rejection) criterion.

The choice of the test statistics or decision variables are not unique. Any strictly monotone function h of a test statistic $\hat{p}$ is also an eligible test statistic since $\{\hat{p} \leqslant p^*\}$ is equivalent to $\{h(\hat{p}) \leqslant h(p^*)\}$. For example, for normal distribution and single (lower) specification limit, we may choose

$$\hat{p} = F_X(L;\hat{\theta}) = \Phi(L;\overline{X},s) = \int_{-\infty}^{L} \frac{1}{s\sqrt{2\pi}} e^{-\left(\frac{x-\overline{X}}{s}\right)^2} dx = \int_{\frac{\overline{X}-L}{s}}^{\infty} \phi(z)dz$$
$$= \int_{Q_L}^{\infty} \phi(z)dz, \tag{3}$$

where $\overline{X}$ and s are the usual sample mean and standard deviation, as the test statistic. This method is known as the *M*-method in Schilling and Neubauer (2009) but we prefer to call it the $\hat{p}$-method in this paper. It is not difficult to see its equivalence to the most popular k-method where Q_L is the test statistic, if the strictly monotonic relation between $\hat{p}$ and Q_L is noticed. However, this equivalence between $\hat{p}$ and Q_L is not valid if the underlying distribution F_X has shape parameters since $\hat{p}$ contains knowledge of shape parameters but Q_L does not.

Therefore, for acceptance sampling procedures under general distributions, in particular those with shape parameters, the $\hat{p}$-method is preferred.

2.3. OC curve

The operating characteristic function (curve) plays central role in the design and implementation of acceptance sampling plans. It measures the performance of a particular sampling plan and is also the basis of choosing target sampling plans which satisfies specific risk management requirements. The OC curve is defined as a function giving the probability of lot acceptance P_a for a given proportion nonconforming p. In the case of a single sampling plan, the acceptance probability is given by

$$P_a = \Pr(\hat{p} \leqslant p^* | p), \tag{4}$$

and this expression implies that the behavior of the OC curve is governed by the distribution of the test statistic $\hat{p}$ conditional on p.