CONSUMER PROTECTION

Surveillance policies on weighing and measuring instruments

GIUSEPPE ARDIMENTO
Weights and Measures Officer
Camera di Commercio Industria Artigianato e Agricoltura di Napoli

EMILIO CLEMENTE
Chief Executive Weights and Measures Officer
Camera di Commercio Industria Artigianato e Agricoltura di Genova

1 Introduction

The surveillance policy operated by a Local Metrology Authority (LMA) is the means by which both consumer protection and the fairness of commercial transactions are ensured in the marketplace.

Generally, the surveillance policy provides for periodical inspection and/or random surveillance to be made on instruments in order to ascertain whether they maintain a steady level of performance within the required accuracy limits during a stated period fixed by law.

The choice of an appropriate weighing and measuring instrument inspection and surveillance policy is nowadays a major concern for LMAs because the new approach *European Directive on Measuring Instruments* (MID) [1] is oriented to conformity assessment by means of procedural modules that predominantly allow manufacturers having sound quality assurance systems to "self-certify" their own products.

Thus, since article 14 of the MID provides for market surveillance operated by EU Member States, the inspection of instruments put into use may be regarded as the only truly independent control in the lifetime of certain instruments [2].

However, in many jurisdictions the large number of devices subjected to legal control does not allow LMAs to efficiently operate in order to ensure fairness in the marketplace and to protect consumers. In these cases it is necessary to entrust the task of periodical inspection to private organizations or laboratories, provided that the latter can guarantee an adequate quality management system for inspections as well as independence and competence in performing inspections, whilst allowing the LMAs to continue their task of carrying out an *a posteriori* sample control on inspected devices in order to monitor the overall performance of those organizations.

2 Main requirements for licensing private verifying organizations or laboratories

A solution which could be implemented in order to solve the problem of assessing the prerequisites of an organization seeking verification licensing is to apply the ISO/IEC 17025 *General requirements for the competence of testing and calibration laboratories* [3].

Indeed legal metrology legislations actually tend to transfer those principles into the national regulation framework, but they often do not provide for an adequate *a posteriori* surveillance; moreover, since the number of weighing and measuring devices used for legal purposes can range from hundreds up to several thousands in many jurisdictions, it is necessary to adopt statistical decision criteria [8] to ascertain whether licensed organizations or laboratories do assure adequate performance when inspecting and certifying instruments.

This paper deals with the attempt to conceive a statistical test method to assess the above quoted performances by means of drawing by the LMAs samples from licensee inspected instrument population and infer a decision based on a stated significance level (see [6] and [7]).

3 The French statistical decision test model

The French regulation *Arrêté du 22 mars 1993 relatif au contrôle des instruments de pesage a fonctionnement non automatique, en service* [4] (*Decree of 22 March 1993 relating to the inspection of nonautomatic weighing instruments in service*) could be of inspiration to design sound statistical decision tests.

In that regulation (articles 10 and 11) criteria on the significance level of statistical tests as well as the minimum licensee verified instrument population size are set out, namely:

- An *a posteriori* control which can be exerted on, say, an annual basis by means of statistical tests must have a *significance level* (see [5] and [6]) of 0.05 or less;

- Licensed laboratories or organizations must perform at least 500 inspections per year in a given LMA jurisdiction;
- The verifying license must be repealed if the *a posteriori* control reveals that the licensee improperly accepts or refuses instruments in a proportion greater than 5 %.

4 The probability model

In order to study the statistical significance level of the decision test to be performed, the binomial distribution can be used: in this instance, the population size has to be considered in respect of the sample size; thus, indicating with:

p — *the probability of finding an instrument <u>properly</u> accepted or refused;*

$q = 1 - p$ — *the probability of finding an instrument <u>improperly</u> accepted or refused.*

(*Note:* p and q are empirically defined as the proportion of properly and improperly accepted or refused instruments in the population subjected to investigation).

The probability of finding exactly r instruments properly accepted or refused in sample of n is:

$$P(r) = \binom{n}{r} \cdot p^{r} \cdot q^{n-r} \qquad (1)$$

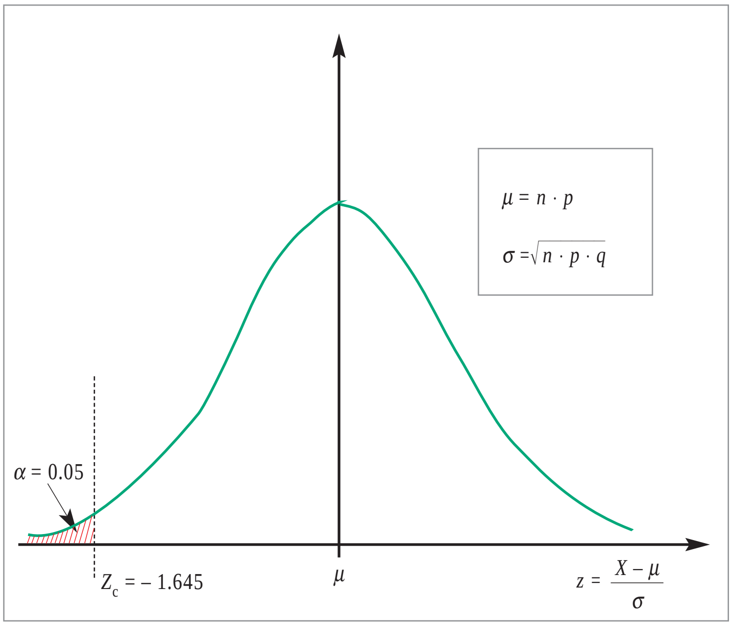

Fig. 1 The normal distribution and the confidence level α along with its one way critical z-value

When large samples are involved in the investigation it becomes more convenient to use a normal distribution, since that is the limit to which the binomial distribution tends as n increases.

The approximation of the binomial distribution by means of the normal distribution can generally be considered as satisfactory where the two following constraints hold simultaneously:

$$n \cdot p \geq 5 \quad \text{and} \quad n \cdot q \geq 5 \qquad (2)$$

If the conditions set out in (2) hold then the normal approximation can be used, thus facilitating calculations. A normal distributed stochastic variable z with zero mean and unity standard deviation can be defined [5] as:

$$z = (X - n \cdot p)/\sqrt{(npq)} \qquad (3)$$

where X is the current number of weighing and measuring instruments properly accepted or refused. In equation (3) use has been made of the above quoted approximation because the population mean (μ) and the standard deviation (σ) characterizing the normal distribution have been set to be equal as follows:

$$\mu = n \cdot p \text{ and } \sigma = \sqrt{(npq)} \qquad (4)$$

(see [5], Chapter 7).

5 The statistical decision test in detail

The statistical decision test is based on the analysis of the z variable as defined in equation (3).

The decision as to whether the licensee verifying performance can be deemed satisfactory shall be taken by minimizing the *I type error* ([5] and [6]), i.e. the probability of rejecting the hypothesis H_0, where H_0 denotes the fact that a licensee performs its work well, for example, properly accepting or refusing at least 95 % of the verified instrument population.

Usually for legal purposes the probability of a *I Type error* (denoted as α) is set to 0.05 as in the French regulation described in paragraph 2.

In Figure 1 a typical one-way decision test is shown: in such a type of statistical decision test the critical z-value (z_c) is $z_c = -1.645$ for the significance level set to $\alpha = 0.05$ (see [5], Chapter 10 and [6]).

Moreover, in the case of poor licensee performance, further investigations are needed in order to evaluate the probability of a *II Type error* [7] associated with the sample based statistical decision.

An acceptable criterion to render the decision reliable with respect to the occurrence of *II Type error* could be the following:

- *The probability (indicated as β) of accepting H_0 hypothesis (p = 0.95) when the actual p equals 0.90 must be less than 0.10.*

An administrative provision of the LMA limiting or even withdrawing the private organizations or laboratories verifying license should always be based on the analysis of the occurrence of the *II Type error* in addition to the *I Type error* one [8].

6 Worked out examples of statistical evaluation of licensee performances

This section deals with the attempt to provide some examples in order to better explain the calculations that should be carried out for evaluating the performances of the licensed verifying laboratories and organizations.

6.1 Example No. 1

A licensed laboratory notifies the LMA of a list of 1680 verified small capacity weighing instruments.

The LMA decides to inspect 150 weighing instruments and finds 142 conforming to the relevant metrological regulations from this sample: in symbols

Population:	N = 1680
Sample size	n = 150
Number of conforming instruments	X = 142

In order to assess the licensee performances,

p = 0.95
$q = 1 - p = 0.05$

and so the expected number of conforming instruments in a sample and the standard deviation are respectively

$\mu = n \cdot p = 142.5$

$\sigma = \sqrt{(n \cdot p \cdot q)} = 2.67$

The z critical value for the test is

$Z_c = -1.645$

From the definition of the z variable in (3), the acceptance criteria can be written as

$$[(X - n \cdot p)/\sqrt{(npq)}] \geq z_c \qquad (5)$$

From equation (5) the acceptance criterion can be written as follows

$$X \geq n \cdot p + z_c \cdot [\sqrt{(npq)}] =$$
$$= n \cdot p - 1.645 \cdot [\sqrt{(npq)}] =$$
$$= X_{LIM} \qquad (6)$$

Since X = 142 and from (6) X_{LIM} = 138, the laboratory passes the *a posteriori* control at a significance level of 0.05.

6.2 Example No. 2

A servicing company is accredited to perform official inspections after a fuel dispenser repair.

In one year it submits to the LMA 1950 "self verification" reports. An *a posteriori* control of the LMA on a sample of 150 reveals that 131 dispensers can be deemed as conforming to the relevant regulations.

In symbols,

N = 1950
n = 150
X = 131

Using the same symbols as in the example 6.1 above, we have

$\mu = n \cdot p = 142.5$

$\sigma = \sqrt{(n \cdot p \cdot q)} = 2.67$

$X_{LIM} = n \cdot p - 1.645 \cdot \sqrt{(npq)} = 142.5 - 1.645 \cdot 2.67 = 138$

Since $X < X_{LIM}$, the acceptance criterion does not hold. But nothing can be said about the performance level delivered by the licensee because only *I Type error* has been investigated: i.e. only the probability of rejecting hypothesis H_0 is below the significance level 0.05.

It should be necessary, in order to decide whether the company performance is poor, to investigate the *II Type error* by using the criterion set out in Paragraph 4:

- *The probability (indicated as β) of accepting H_0 hypothesis (p = 0.95) when the actual p equals 0.90 must be less than 0.10.*

That graphically corresponds to the situation depicted in Figure 2.

In order to reasonably state that the company performances are not satisfactory to the LMA granting the verifying license, the area β (the so-called *consumer risk* [5]) should be less than or equal to 0.10

Thus, in order to evaluate β let

p = 0.90

$q = 1 - p = 0.10$

The corresponding values for μ and σ are

$\mu_\beta = n \cdot p = 150 \cdot 0.90 = 135$

$\sigma_\beta = \sqrt{(n \cdot p \cdot q)} = 3.67$

Thus

$z_\beta = (X_{LIM} - \mu_\beta)/\sigma_\beta = (138 - 135)/3.67 = 0.82$

$\beta = Probability\ (z \geq z_\beta) = 0.21$

The criterion set out in paragraph 4 and above is not met; thus the sample should be enlarged to render the decision based on the sample test more reliable.

A further 50 dispensers are then randomly drawn from the population and inspected by the LMA. The results it achieves, considering the whole sample, are

$n = 200$

$X = 170$

Let

p = 0.95

$q = 1 - p = 0.05$

We then have

$\mu = n \cdot p = 200 \cdot 0.95 = 190$

$\sigma = \sqrt{(n \cdot p \cdot q)} = 3.08$

$X_{LIM} = \mu - z_c \cdot \sigma = 190 - 1.645 \cdot 3.08 = 184$

Since $X < X_{LIM}$ the company does not pass the *I Type error* criterion. In order to establish whether the decision is well-founded that the company performances are not satisfactory, let us analyze the *II Type error.*

Let

$p = 0.90$

$q = 1 - p = 0.10$

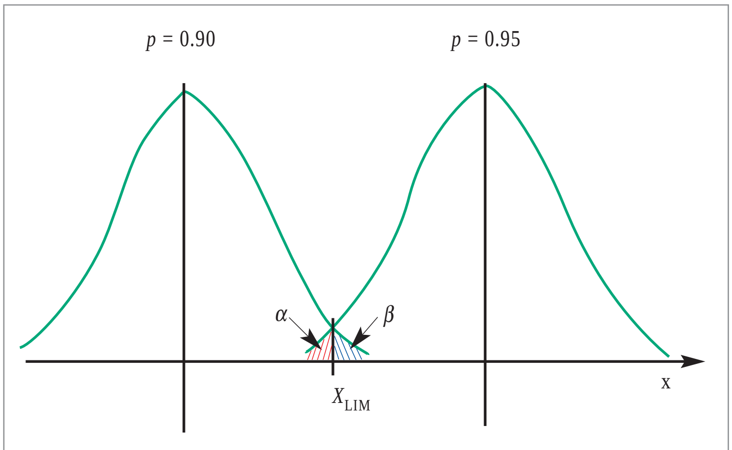

Fig. 2 α-confidence level, β-confidence level and the probability density for good performance (p =0.95) and for poor performance (p = 0.90)

We have

$\mu_\beta = n \cdot p = 200 \cdot 0.90 = 180$

$\sigma_\beta = \sqrt{(n \cdot p \cdot q)} = 4.24$

Thus,

$z_\beta = (X_{LIM} - \mu_\beta)/\sigma_\beta = (184 - 180)/4.24 = 0.94$

$\beta = Probability\ (z \geq z_\beta) = 0.17$

Also this time the *II Type error* criterion is not met: thus we need to further increase the sample size in order to achieve the required significance level. A further sub-sample of 50 dispensers is drawn from the population of licensee verified instruments.

The overall results obtained by the LMA are

$n = 250$

$X = 214$

Let

p = 0.95

$q = 1 - p = 0.05$

We have then

$\mu = n \cdot p = 250 \cdot 0.95 = 237.5$

$\sigma = \sqrt{(n \cdot p \cdot q)} = 3.45$

$X_{LIM} = \mu - z_c \cdot \sigma = 237.5 - 1.645 \cdot 3.45 = 231$

Since $X < X_{LIM}$ the LMA confirms that the company does not pass the *I Type error* criterion. The analysis of the *II Type error* yields:

Let

$p = 0.90$

$q = 1 - p = 0.10$

We have

$\mu_\beta = n \cdot p = 250 \cdot 0.90 = 225$

$\sigma_\beta = \sqrt{(n \cdot p \cdot q)} = 4.74$

Thus,

$z_\beta = (X_{LIM} - \mu_\beta)/\sigma_\beta = (231 - 225)/4.74 = 1.27$

$\beta = Probability\ (z \geq z_\beta) = 0.10$

Based on the results obtained the LMA may reasonably conclude that the servicing performance of the company is not satisfactory and thus the quality assurance system on which the verifying license was granted must be reviewed; meanwhile the LMA will have to limit, suspend or even withdraw the license.

7 Acknowledgements

The authors wish to thank Mr Jean-François Magaña (BIML Director); they also extend their thanks to Mrs. Corinne Lagauterie (Sous-Direction de la Métrologie, France) for supplying useful information about Reference 4, Mr Constantine V. Cotsoradis (Kansas Department of Agriculture) who gave detailed information about the statistical testing program on the performance of licensed servicing companies operating in the state of Kansas, USA and, last but not least, Dr. Walter Bich (Istituto di Metrologia G. Colonnetti, Italy) as well as Prof. Giulio Barbato (Politecnico di Torino, Italy) for their courtesy in granting valuable hints and comments to this paper.

References

[1] Directive of the European Parliament and of the Council on Measuring instruments (MID), Commission proposal 15/09/2000

[2] Michael Harvey, *The Balance must swing towards inspection*, The Trading Standards Review, October 1998

[3] ISO – IEC 17025: *General requirements for the competence of testing and calibration laboratories*

[4] Arrêté du 22 mars 1993 relatif au contrôle des instruments de pesage a fonctionnement non automatique, en service, J.O. de la République Française du 28 mars 1993

[5] Murray R. Spiegel, *Statistics*, Mc Graw-Hill

[6] ISO 2854: *Statistical interpretation of data - Techniques of estimation and tests relating to means and variables*

[7] ISO 3494: *Statistical interpretation of data - Power tests relating to means and variances*

[8] Nello Polese, *Misure per la gestione*, Collana di Ingegneria Economico - Gestionale, Edizioni Scientifiche Italiane

GIUSEPPE ARDIMENTO

Weights and Measures Officer

Camera di Commercio Industria Artigianato e Agricoltura di Napoli

EMILIO CLEMENTE

Chief Executive Weights and Measures Officer

Camera di Commercio Industria Artigianato e Agricoltura di Genova