WEIGHING

Verification of weighing instruments from a statistical point of view

Teuvo Lammi, The Finnish Association of Technology for Weighing, Helsinki, Finland

1 Introduction

This paper deals with the verification of nonautomatic, single interval weighing instruments from a statistical point of view.

On the basis of the verification test results obtained for weighing instruments, the verification officer makes the decision as to whether or not an instrument can be verified.

The test results are estimates, i.e. their values are associated with uncertainties and due to them the officer may make incorrect decisions.

The aim of this paper is to investigate these decisions and to make suggestions about how to avoid them. A formula is given for the uncertainty of the errors in the indication of the instrument observed in the weighing test (R 76-1, A.4.4.1). It is used in the study of incorrect decisions and also to judge some of the requirements laid down for verification.

In Section 2 a short note on the verification tests is given. Sections 3 and 4 deal with incorrect decisions. In Section 5 a formula for the uncertainty associated with the results of the weighing test is presented.

2 Notes on verification tests

The flow chart at the bottom of this page shows some of the verification tests and checks for the instruments.

1)	2)	3)
Checks before the tests, e.g. leveling, connection to the power supply and temperature stability	Weighing test where the indications are compared with the values of the test loads (standard weights) - the “true” values of the indications	Repeatability and eccentricity tests where differences between the results of several weighings of the same load are investigated

The test results in 2) must be within the MPEs, *the maximum permissible errors on initial verification* (R 76-1, 3.5), and the differences between the results of the weighings in 3) must meet the permissible differences (R 76-1, 3.6).

3 Uncertainty and a “quality” indicator for verification

3.1 “True” E

For a certain load let E be the error of the instrument obtained in the weighing test and U the value of the uncertainty of that error. The interval E ± U covers the “true” value of E with a “high” confidence level. The “true” value of E is here called the “true” E.

According to the requirements of R 76-1, 3.5 the absolute value of the error E must satisfy the condition

$$|E| \le |MPE|$$

for all the loads. The question is, what is the probability that

$$|\text{“true” } E| \le |MPE|$$

is true when

$$|E| \le |MPE|$$

is met and U takes on different values?

3.2 Probability that |“true” E| ≤ |MPE| is true *)

Case 1: U ≤ 1/3 × |MPE|

If $|E| \le 2/3 \times |MPE|$ and $U \le 1/3 \times |MPE|$, then substituting these values for E and U in E ± U (which includes the “true” E) it is easy to see that $|\text{“true” } E| \le |MPE|$ is true.

---

*) A similar discussion of this subject is given in the author's paper “Calibration of Weighing Instruments and Uncertainty of Calibration”, OIML Bulletin, October 2001.

In general, if $|E| \leq |MPE|$ and $U \leq 1/3 \times |MPE|$, the probability P that $|$"true" $E| \leq |MPE|$ is true is approximated by the fraction

$$|MPE| / (|MPE| + 1/3 \times |MPE|).$$

Now $|MPE|$ is half the length of the interval where the "true" E should be and

$$|MPE| + 1/3 \times |MPE|$$

that where it is. If $U < 1/3 \times |MPE|$, P is greater than the fraction and if $U = 1/3 \times |MPE|$, P equals the fraction. So:

$$P \geq |MPE| / (|MPE| + 1/3 \times |MPE|) =$$
$$= |MPE| / (4/3 \times |MPE|) = 75\ \%$$

Case 2: U < |MPE|

$U = k \times |MPE|$ $(k < 1)$. In a similar way as in Case 1 the probability P that $|$"true" $E| \leq |MPE|$ is true is:

$$P = |MPE| / (|MPE| + k \times |MPE|) =$$
$$= 1 / (1 + k) > 50\ \%\ (k < 1)$$

Example:

Let the observed E be

$$E = +\ 0.4 \times |MPE|.$$

If $k = 0.9$, then the "true" E is in the interval

$$E \pm 0.9 \times |MPE| \text{ (its length is } 1.8 \times |MPE|).$$

In order for the condition

$$|\text{"true" } E| \leq |MPE|$$

to be true, the "true" E should be in the interval

$$\text{from } -0.5 \times |MPE| \text{ to } |MPE|$$

the length of which is $1.5 \times |MPE|$. Thus $P = 1.5 \times |MPE| / (1.8 \times |MPE|) \approx 83\ \%$.

Case 3: U ≥ |MPE|

$U = k \times |MPE|$ $(k \geq 1)$. The probability P that $|$"true" $E| \leq$ MPE is true is:

$$P = |MPE| / (|MPE| + k \times |MPE|) =$$
$$= 1 / (1 + k) \leq 50\ \%\ (k \geq 1)$$

On the basis of the previous cases one can draw the conclusion that the smaller the value U assumes, the better the chances are that $|$"true" $E| \leq |MPE|$ is true when $|E| \leq |MPE|$.

3.3 "Quality" indicator U

If $U < |MPE|$ $(P > 50\ \%)$, the quality of the verification is here regarded as good enough. Obviously values of

$$U \leq 1/3 \times |MPE| \ (P \geq 75\ \%)$$

are ideal but may sometimes be difficult to achieve. Practical conditions for $U < |MPE|$ are given in 5.3.2 and for $U \leq 1/3 \times |MPE|$ in 5.3.3.

If $U \geq |MPE|$ $(P \leq 50\ \%)$, the values of U should be reduced by having the instrument serviced and adjusted. As stated in 5.3.2 the adjustment should primarily aim to reduce the eccentric errors and the repeatability error, if possible. The intention is: $U < |MPE|$.

4 Type I and II errors and OC-curves

4.1 Type I and II errors

Consider "Type I" in Figure 1 where the observed E (3.1) is E > +MPE. If the "true" E in the interval E ± U is "true" E < +MPE, it complies with the requirements (a "good" result). However, the observed E is E > +MPE and does not comply with the requirements. Because E is the basis for decision, a Type I error is committed (the "good" result cannot accepted).

Consider "Type II" in Figure 1 where E < +MPE. If the "true" E in the interval E ± U is "true" E > +MPE, it does not comply with the requirements (a "poor" result). However, the observed E is E < +MPE and complies with the requirements. Because E is the basis for decision, a Type II error is committed (the "poor" result is accepted).

Type I and II errors can also be brought about by some defects in the tests (Section 2). For example:

A) If in the eccentricity test the variations in the zero point are not taken into account accurately enough before the test load is applied to the different positions on the load receptor, then the results of the test may be misleading and the decisions made on their basis may be incorrect.

B) Suppose that the errors of the indications obtained in the weighing test vary in a non-linear way and that they are within the MPEs. However, the errors of the net values may exceed the MPEs. If in this case the errors of the net values are not investigated as they should be, an instrument not complying with the requirements might be verified and a Type II error is committed.

4.2 OC-curves

In the following the effect of Type I and II errors is illustrated with the aid of OC-curves (see textbooks dealing with statistical quality control) showing the probability that the instrument is verified.

4.2.1 Ideal OC-curve

Let us deal with an imaginary case where E is within the MPEs but U equals zero. Thus the observed E equals the "true" E. It is thus possible to perform the verification without the effect of Type I and II errors. This is illustrated by the ideal OC-curve in Figure 2.

4.2.2 Actual OC-curve

In real situations the uncertainty associated with the observed E differs from zero. When this E is used to investigate whether or not the condition $|E| \leq |MPE|$ (3.1) is met, incorrect decisions can be made due to Type I and II errors as explained in 4.1.

Consider Figure 3 where two example OC-curves are shown. Their ordinates show the probability P that the instrument is verified. Now let a Type I error mean that a "good" instrument (all the "true" E values are within the MPEs) is not verified and a Type II error that a "poor" instrument (all the "true" E values are not within the MPEs) is verified.

In order to avoid these errors P should be as large as possible when |"true" E | ≤ |MPE| and as small as possible when |"true" E | > |MPE|.

Curve a) in Figure 3

Type I errors may be committed because P < 1 for the values of |"true" E | which are just below |MPE |. So "good" instruments may sometimes not be verified. If the |"true" E | is "small" or near zero, then P ≈ 1 and Type I errors can very likely be avoided.

Type II errors can be committed because P > 0 for the values of |"true" E | which are slightly greater than |MPE |. So a "poor" instrument may be verified, although in this case quite rarely. P decreases as |"true" E | increases and assumes zero if |"true" E | is great enough. So the chances of Type II errors gradually decrease as |"true" E | increases.

Curve a) is considered to be a good fit to the step curve (the ideal OC-curve). The fit is better the smaller the values U assumes. On the other hand, the better the fit the more unlikely Type I and Type II errors are.

Curve b) in Figure 3

For the values of | "true" E | which are slightly smaller than |MPE |, P assumes values zero. So Type I errors are very likely and "good" instruments are in practice not verified. However, if | "true" E | is near zero, P ≈ 1 and the very "good" instruments (|"true" E | ≈ 0) can be verified.

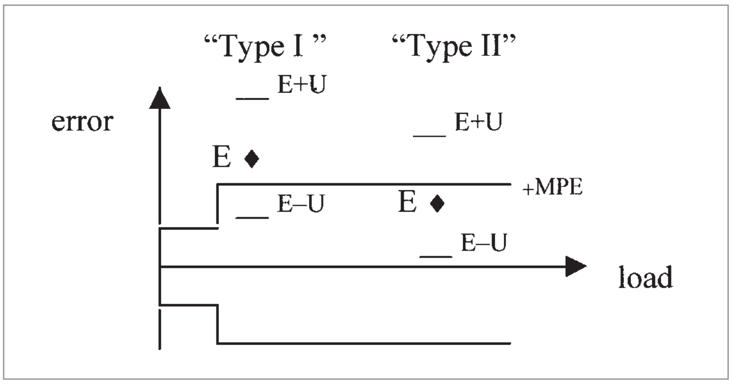

Figure 1

"Type I": Let the "true" E be < +MPE. Decisions are made according to the observed E which is E > +MPE. So a Type I error is committed.

"Type II": Let the "true" E be > +MPE. Decisions are made according to the observed error E which is E < +MPE. So a Type II error is committed.

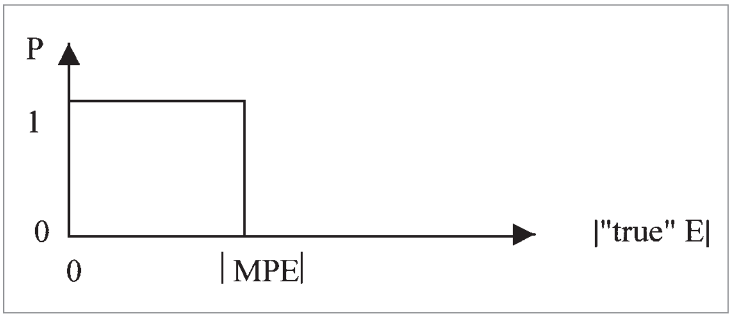

Figure 2

If | "true" E| ≤ | MPE | (U = 0), the probability P that the instrument is verified is 1.

If | "true" E| > | MPE|, the probability is 0.

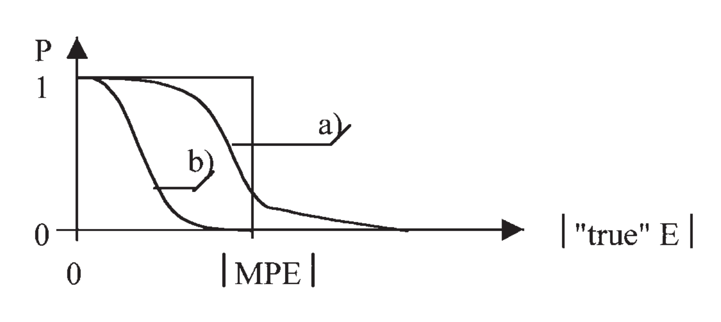

Figure 3

If |"true" E| ≤ |MPE| (U > 0), the probability P (curve a) that the instrument is verified is ≤ 1. Only if the values of |"true" E| are near zero, then P = 1.

If |"true" E| > |MPE|, P (curve a) is > 0 and cannot be 0 until the values of |"true" E| are great enough. The fit of curve a) to the step curve (the ideal OC-curve) is quite good but that of curve b) is not.

Type II errors are practically impossible and "poor" instruments are not likely to be verified at all. This is achieved at the expense of committing Type I errors.

Curve b) could represent a situation where instead of |E| ≤ |MPE| the requirement |E ± U| ≤ |MPE| is applied to verification. The fit of curve b) to the step curve is considered to be very poor.

5 Practical evaluation of the uncertainty and requirements

5.1 Formula for U

The uncertainty U associated with the errors E (3.1) obtained in the weighing test is evaluated here with the aid of the following formula for U:

$$U = 2r\,[(k_n\,R)^2 + u^2 + (0.4\,\Delta)^2]^{1/2}$$

Where: *)

R is the repeatability, i.e., the difference between the largest and the smallest results in the repeatability test. The test load is the largest load used in the test. Frequently, it is near Max (R 76-1. A.4.10).

k_n R is the standard deviation of the results of the repeatability test. k_n assumes the following values according to the number n (n ≥ 3) of results in the test:
$k_3 = 0.591$, $k_4 = 0.486$, $k_5 = 0.430$, $k_6 = 0.395$,
$k_7 = 0.370$, $k_8 = 0.350$, $k_9 = 0.337$, $k_{10} = 0.325$.

u is the standard deviation of the errors of the verified weights used. u = 0.4 × (the sum of the |mpe|s of the weights for the load which corresponds to the load used in the repeatability test).

Δ is the greatest eccentric error noted in the eccentricity test (R 76-1, A.4.7). Frequently, the test load is 1/3 × |MPE| of the instrument. If Δ is less than or equal to the smaller of |Δ| < |MPE| or |Δ| < e for the load used in the test, set Δ = 0 in U. In this case the errors in the weighing test can be regarded as independent of the positions of the weights on the load receptor. Otherwise, Δ ≠ 0 and 0.4|Δ| is the standard deviation of the errors brought about by the eccentric positions of the weights on the load receptor during the weighing test.

r is a coefficient and assumes the values 0.3, 0.4, 0.7 and 1 which are associated with the values of the MPEs of the instrument as given in Table 1. r is used to evaluate U for the loads where the MPEs take on the different values ± 0.5 e, ± 1 e and ± 1.5 e **or** ± 0.5 e and ± 1 e **or** only ± 0.5 e.

The formula for U can be used if:

A) digital rounding errors included in digital indications are eliminated (R 76-1, 3.5.3.2),

B) readings of the indications are unambiguous (R 76-1, 4.2.1),

C) the verification is performed at a steady ambient temperature (R 76-1, A.4.1.2),

D) verified weights are used in the verification, and

E) the buoyancy effect of the air density on weights does not need to be taken into account (note that this effect should also be considered on the load measuring device (load cell) and the load receptor).

5.2 Determination of U

The values of k_n R, u and Δ are determined as mentioned in 5.1 and are inserted in the formula for U. Thereafter, according to r (Table 1) the values of U are sequentially evaluated for the loads where the MPEs take on the different values.

*) Explanations of k_n R, u, Δ and r are presented in the author's paper "Calibration of Weighing Instruments and Uncertainty of Calibration", OIML Bulletin, October 2001.

For example, let the instrument be of class III and Max/e = n = 2000. Thus the MPEs assume the values ± 0.5 e and ± 1 e. U is as follows:

$$U = 2 \times 0.4 \times [(k_n\ R)^2 + u^2 + (0.4\ \Delta)^2]^{1/2}$$

for the loads where

$$MPE = \pm 0.5\ e\ (r = 0.4)$$

$U = 2 \times [(k_n\ R)^2 + u^2 + (0.4\ \Delta)^2]^{1/2}$ for the loads where

$$MPE = \pm 1\ e\ (r = 1).$$

5.3 Requirements and values of U

5.3.1 Values of U expressed in terms of e

Let us deal with instruments with MPEs which take on the values ± 0.5 e and ± 1 e.

According to the requirements of R 76-1 the values of R, u and Δ could be as follows:

- R can be at most e, if the test load is near Max. The number of weighings is supposed to be six. So $k_6 = 0.395$ and $k_6\ R \approx 0.4\ e$ (R 76-1, 3.6.1 and 8.3.3).
- In order to obtain u, calculate the sum of the |mpe|s of the weights Σ|mpe| for the test load used for R. So according to u in 5.1 and R 76-1, 3.7.1, $u = 0.4 \times \Sigma|mpe| \leq 0.4 \times 1/3 \times |MPE| = 0.4 \times 1/3\ e$ because |MPE| = e of the instrument for the load in question.
- The value of |Δ| can be at most e (R 76-1, 3.6.2).

Table 1 Coefficient r

The values of r associated with the values of the MPEs for the instrument.			
MPE:	±0.5 e	± 1 e	± 1.5 e
r:	0.3	0.7	1
r:	0.4	1	-
r:	1	-	-

Insert these greatest values for k_n R, u and Δ in U. Thus, the value of U for the loads for which MPE = ± 0.5 e (r = 0.4) is:

$$U = 2r[(k_n\ R)^2 + u^2 + (0.4\ \Delta)^2]^{1/2} =$$
$$= 2 \times 0.4 \times [(0.4\ e)^2 + (0.4 \times 1/3\ e)^2 + (0.4\ e)^2]^{1/2} =$$
$$= 2 \times 0.4 \times 0.58\ e \approx 0.46\ e\ (U \approx |MPE|)$$

The value of U for the loads for which MPE = ±1 e (r = 1) is:

$$U = 2r[(k_n\ R)^2 + u^2 + (0.4\ \Delta)^2]^{1/2} =$$
$$= 2 \times [(0.4\ e)^2 + (0.4 \times 1/3\ e)^2 + (0.4\ e)^2]^{1/2}$$
$$= 2 \times 0.58\ e \approx 1.2\ e\ (U > |MPE|)$$

In a similar way the values of U can be approximated if the MPEs of the instrument assume the values ± 0.5 e, ± 1 e and ± 1.5 e or only ± 0.5 e.

5.3.2 Conditions for U < |MPE| and suggestions for R, Δ and errors of the weights

In order to arrive at values of U which are smaller than |MPE| (see 3.2 and 3.3), the following values are suggested for R, Δ and the errors of the weights:

- R should be R < |MPE| or R < e for the applied test load, whichever is smaller. The number n of weighings in the repeatability test should be n ≥ 5 (the values of k_n (5.1) are quite stable for these values of n and thus the information from the test could be good enough).
- |Δ| should be |Δ| < |MPE| or |Δ| < e for the applied test load, whichever is smaller. In this case set Δ = 0 in U.
- The weights for the weighing test should, if possible, be selected so that their errors are not greater than 1/5 (instead of 1/3) of the |MPE| of the instrument for the applied load.

5.3.3 Conditions for U ≤ 1/3 × |MPE|

If U should be U ≤ 1/3 × |MPE| and the MPEs assume the values ± 0.5 e and ± 1 e, then R should be R ≤ 0.35 e for the applied test load while Δ is as given in 5.3.2. The weights should preferably be selected so that their errors are at most 1/5 (instead of 1/3) of the |MPE| for the applied load.

If U should be U ≤ 1/3 × |MPE| and the MPEs take on the values ± 0.5 e, ± 1 e and ± 1.5 e, R should be R < 0.55 for the applied test load, Δ is as given in 5.3.2 and 1/5 should be used in the selection of the weights. However, if the weights are selected using 1/3, R should be R < 0.4 e for the applied test load while Δ is as given in 5.3.2. ■

References

OIML Recommendation R 76-1: *Non-automatic weighing instruments. Part 1: Metrological and technical requirements - Tests* (1992)

Guide to the Expression of Uncertainty in Measurement, BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML (1995 corrected and reprinted edition)