DUBAI ACCREDITATION CENTER – DUBAI MUNICIPALITY

GUIDANCE FOR STATISTICAL EVALUATION OF INTERLABORATORY PROFICIENCY TESTING PROGRAM

DAC-G3-03

June 2008

CONTENTS

Forward 3
1. Objective 4
2. Scope 4
3. Homogeneity Check of Samples 4
4. Stability Check of Samples 4
5. Evaluation of Results 4
6. Example 6
7. References 10

Forward

Dubai is a rapidly expanding Emirate, and the Government places great emphasis on providing quality services. The main role of DM is to formulate the Emirate's urban strategic plans as well as the provision of essential infrastructure, environmental and health services for the continued development of Dubai as a modern, safe and dynamic Emirate. The DM, through its Dubai Accreditation Department (DAC) undertakes assessment and accreditation of various Conformity Assessment Bodies (CABs) according to International Standards, guidelines and world best practices.

The Dubai Accreditation Department, hereinafter referred to as the DAC, is the sole authority for the granting, maintenance, suspension or withdrawal of accreditation.

It is the responsibility of the DAC to ensure that the latest version of this document is available for reference. The copyright of this text is held by DAC, copies cannot be sold and DAC publications are not allowed to be photocopied or electronically reproduced by DAC assessors and accredited CABs unless permission is given by DAC, the Printed copy of DAC publications documents shall be treated as 'Uncontrolled'. Always refer to the controlled version available on DAC web site.

For further information about this publication, please check our website for up-to-date information www.dac.gov.ae

1. Objectives

To define DAC policy in organizing, conducting and analyzing the Inter-Laboratory Proficiency Testing Program (PTP) based on ISO 13528 Standard.

2. Scope

This guidance document is being issued for the purpose of providing a brief description of the statistical method used in determining the unacceptable results of inter laboratory proficiency testing program. and the statistical methods used in determining the homogeneity and stability check and z-score of inter-laboratory proficiency testing program conducted by Dubai Accreditation Department (DAC). Z-score result is used to identify outlier laboratory.

3. Homogeneity Check of Samples

3.1 It is analytically verified and ensured that all PT samples within a packaging are adequately homogenous and stable (when applicable) prior to their use in a PT program.

3.2 Samples, or parameters, which fail to pass the homogeneity testing criteria, cannot be used in DAC PT program to evaluate its accredited laboratories.

3.3 Certified Reference Materials (CRM) are used in some programs such as environmental and gold PTP where homogeneity is ensured by the supplier.

3.4 Appropriate statistical techniques based on ISO 13528: 2005 are applied for the evaluation of data from homogeneity and stability testing of test items.

4. Stability Check of Samples

4.1 The stability tests should be conducted after the closing date for the round, prior to formal evaluation of participant results.

4.2 Samples shall be stored under conditions similar as those required of participant laboratories.

4.3 The allowed time delay between homogeneity tests and the stability tests shall be similar to the time delay that will be experienced by the samples tested by the participants in the PT.

5. Evaluation of Results

5.1 Assigned Value and Uncertainty

5.1.1 The evaluation of results is done in accordance with the ISO 13528: 2005, using statistical techniques based on robust analysis (algorithm), this yields robust values of the average; standard deviation and z-scores of the data to which it is applied, this is including homogeneity and stability checks. The estimates of robust standard deviation and robust average are derived by an iterative calculation by updating the values several times (iteration 0 - 6) using the

modified data, until the process converges. Convergence may be assumed when there is no change from one iteration to the next in the third significant figure of the robust standard deviation and the equivalent robust average. This statistical process is designed as Excel template (Table – 1)

5.1.2 The assigned value and its uncertainty will be derived after the round is completed. The assigned value X for the test material used in a round of a proficiency testing scheme is the robust average of the results reported by all the participants in the round. The assigned value X is calculated as the robust average of the results reported by the laboratories, using Algorithm A in Annex C (ISO 13528: 2005), as follow:

5.1.3 The robust estimation x* and s* may be derived by an iterative calculation, i.e. by updating the values of x* and s* several times using the modified data, until the process converges. Convergence may be assumed when there is no change from one iteration to the next in the third significant figure of the robust standard deviation and of the equivalent figure in the robust average. These calculations have been reflected in the example (6) given below.

5.2 Use of z-scores

5.2.1 Z-Score, is standardized measure of laboratory bias, calculated using the assigned value and the standard deviation for proficiency assessment. This statistical procedure will use assigned z-scores to identify outlier results.

5.2.2 When a participant reports a result that gives rise to a z-score above +3.0 or below -3.0, then the result shall be considered (an outlier) to give an “action signal”. Likewise, a z-score above +2.0 or below -2.0 shall be considered to give a “warning signal”. A single “action signal”, or “warning signals” in two successive rounds, shall be taken as evidence that an anomaly has occurred that requires investigation.

5.2.3 The calculation of the z-scores depends on the statistical design of the program as explained above.

5.2.4 These robust z-scores will use the results of participants; the median and the New s* as indicated above, any result which has a z-score outside the range ±3 will be identified as an outlier.

5.2.5 The z-score for a result is the result minus the median of iteration 6 divided by the New s* iteration 6, as mentioned above.

5.3 Identifying unacceptable results

A very high (>+3) between-laboratories z-score indicates that a laboratory’s results is significantly higher than the consensus value (median). Similarly, a very low (i.e. negative, <-3) between-laboratories z-score show that a laboratory’s result is lower than expected.

5.4 Graphical presentation of the results

Ordered z-score bar charts will be used to illustrate the data. Outlier laboratories will be located at the extreme ends of the bar chart with their corresponding bars exceeding the ±3 axes. Laboratories with z-scores near the median or the consensus result will be located at or near the center of the chart with their corresponding bars at or near zero.

In addition, the familiar Youden x-y plot is also used in order to show the position of each participating results in relation to each other. The chart will also indicate the presence of systematic errors by the location of a laboratory's data point in the upper right or lower left quadrant.

***Note**: if proficiency testing schemes that involve very large numbers of laboratories (e.g. over 100 laboratories), the normal probability plots may be used to supplement the interpretation of the z-score. But if there are only a small number of laboratories (e.g. less than 10 laboratories), no signal may be given. In this case, the graphical methods that combine scores over several rounds will provide more useful indications of the performance of the laboratories than the results of individual rounds.*

6. Example

6.1 In this example the data from the 155th Inter-laboratory Proficiency Testing Program organized by DAC for (20) participants involving the determination of compressive Strength of Paving Block samples is used. Each sample consisted of 10 specimens, and the average of each ten results was used in the statistical calculations.

6.2 As per experimental design, the results for each participating laboratory are listed in iteration 0. The z-scores and other statistics mentioned in clause [5.2] are calculated and the results are tabulated in the below table.

Table (1) statistical analysis to calculate z-score and assigned the target value in compliance with ISO 13528:2005

Iteration	0	xi-x*	1	$(xi-x^*)^2$	2	$(xi-x^*)^2$	3	$(xi-x^*)^2$	4	$(xi-x^*)^2$	5	$(xi-x^*)^2$	6	$(xi-x^*)^2$	Z Score
δ = 1.5 s*	---		4.45		4.42		4.35		4.32		4.31		4.30		
x* - δ	---		40.55		40.78		40.87		40.91		40.93		40.93		
x* + δ	---		49.45		49.62		49.57		49.55		49.54		49.54		
Lab 16	40	5.00	40.55	21.61	40.78	19.75	40.87	19.00	40.91	18.70	40.93	18.58	40.93	18.53	-1.83
Lab 8	40	5.00	40.55	21.61	40.78	19.75	40.87	19.00	40.91	18.70	40.93	18.58	40.93	18.53	-1.83
Lab 3	41	4.00	41.00	17.64	41.00	17.83	41.00	17.91	41.00	17.94	41.00	17.96	41.00	17.96	-1.48
Lab 6	43	2.00	43.00	4.84	43.00	4.94	43.00	4.98	43.00	5.00	43.00	5.01	43.00	5.01	-0.78
Lab 10	44	1.00	44.00	1.44	44.00	1.50	44.00	1.52	44.00	1.53	44.00	1.53	44.00	1.53	-0.43
Lab 12	44	1.00	44.00	1.44	44.00	1.50	44.00	1.52	44.00	1.53	44.00	1.53	44.00	1.53	-0.43
Lab 19	44	1.00	44.00	1.44	44.00	1.50	44.00	1.52	44.00	1.53	44.00	1.53	44.00	1.53	-0.43
Lab 13	45	0.00	45.00	0.04	45.00	0.05	45.00	0.05	45.00	0.06	45.00	0.06	45.00	0.06	-0.08
Lab 21	45	0.00	45.00	0.04	45.00	0.05	45.00	0.05	45.00	0.06	45.00	0.06	45.00	0.06	-0.08
Lab 4	45	0.00	45.00	0.04	45.00	0.05	45.00	0.05	45.00	0.06	45.00	0.06	45.00	0.06	-0.08
Lab 9	45	0.00	45.00	0.04	45.00	0.05	45.00	0.05	45.00	0.06	45.00	0.06	45.00	0.06	-0.08
Lab 11	46	1.00	46.00	0.64	46.00	0.60	46.00	0.59	46.00	0.58	46.00	0.58	46.00	0.58	0.27
Lab 7	46	1.00	46.00	0.64	46.00	0.60	46.00	0.59	46.00	0.58	46.00	0.58	46.00	0.58	0.27
Lab 15	47	2.00	47.00	3.24	47.00	3.16	47.00	3.13	47.00	3.11	47.00	3.11	47.00	3.10	0.61
Lab 17	47	2.00	47.00	3.24	47.00	3.16	47.00	3.13	47.00	3.11	47.00	3.11	47.00	3.10	0.61
Lab 2	47	2.00	47.00	3.24	47.00	3.16	47.00	3.13	47.00	3.11	47.00	3.11	47.00	3.10	0.61
Lab 5	47	2.00	47.00	3.24	47.00	3.16	47.00	3.13	47.00	3.11	47.00	3.11	47.00	3.10	0.61
Lab 1	48	3.00	48.00	7.84	48.00	7.71	48.00	7.66	48.00	7.64	48.00	7.63	48.00	7.63	0.96
Lab 14	56	11.00	49.45	18.05	49.45	17.86	49.45	17.78	49.45	17.75	49.45	17.74	49.45	17.73	3.75
Lab 18	60	15.00	49.45	18.05	49.45	17.86	49.45	17.78	49.45	17.75	49.45	17.74	49.45	17.73	5.15
Average	46.00		45.20	128.37	45.22	124.24	45.23	122.57	45.24	121.90	45.24	121.63	45.24	121.52	
SD	4.74		2.60	6.76	2.56	6.54	2.54	6.45	2.53	6.42	2.53	6.40	2.53	6.40	
New x*	45	2.00	45.200	2.60	45.223	2.56	45.232	2.54	45.236	2.53	45.24	2.53	45.24	2.53	
New s*	2.97		2.948		2.900		2.880		2.872		2.87		2.87		

Explanation of the calculations in the table above

Guidance for other calculations and identifications of some number in table (2)

N = number of participant labs	20
Average of participant's results	46
Standard deviation of participant's results (SD)	4.74
New **x*** = the median of participant's results	45
New **s*** = 2.00 * 1.483	2.97
The median of the second column ($\mathbf{x_i}$ - **x***)	2.00
Consensus assigned value of the distributed samples	45± 2.87 (robust mean and SD)

Evaluation of participant's results and z-score is as follow:

Target value	45.24 (New **x*** column 6)
Low Acceptable	45.24-3*2.87(New **s*** column 6)=36.63
High Acceptable	45.24+3*2.87(New **s*** column 6)= 53.85
Acceptable Range	36.63 – 53.85

Z-Score is calculated as follow: $\frac{\text{the result (from column 0) - New } \mathbf{x^*} \text{ (column 6)}}{\text{New } \mathbf{s^*} \text{ (column 6)}}$

Lab 14 = $\frac{56 - 45.24}{2.87}$ = 3.753

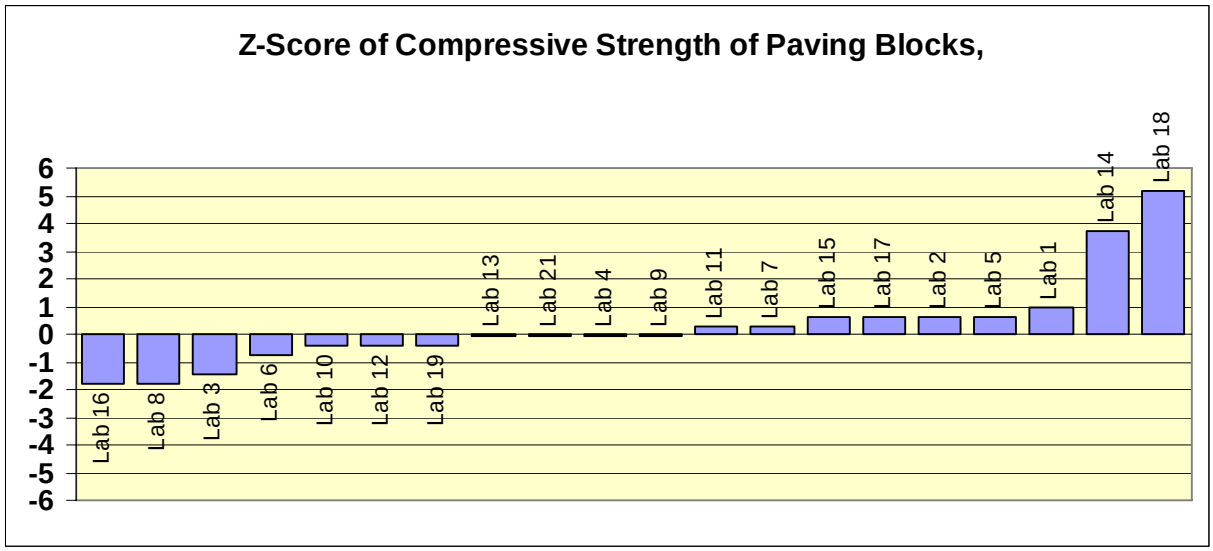
Lab 18 = $\frac{60 - 45.24}{2.87}$ = 5.143

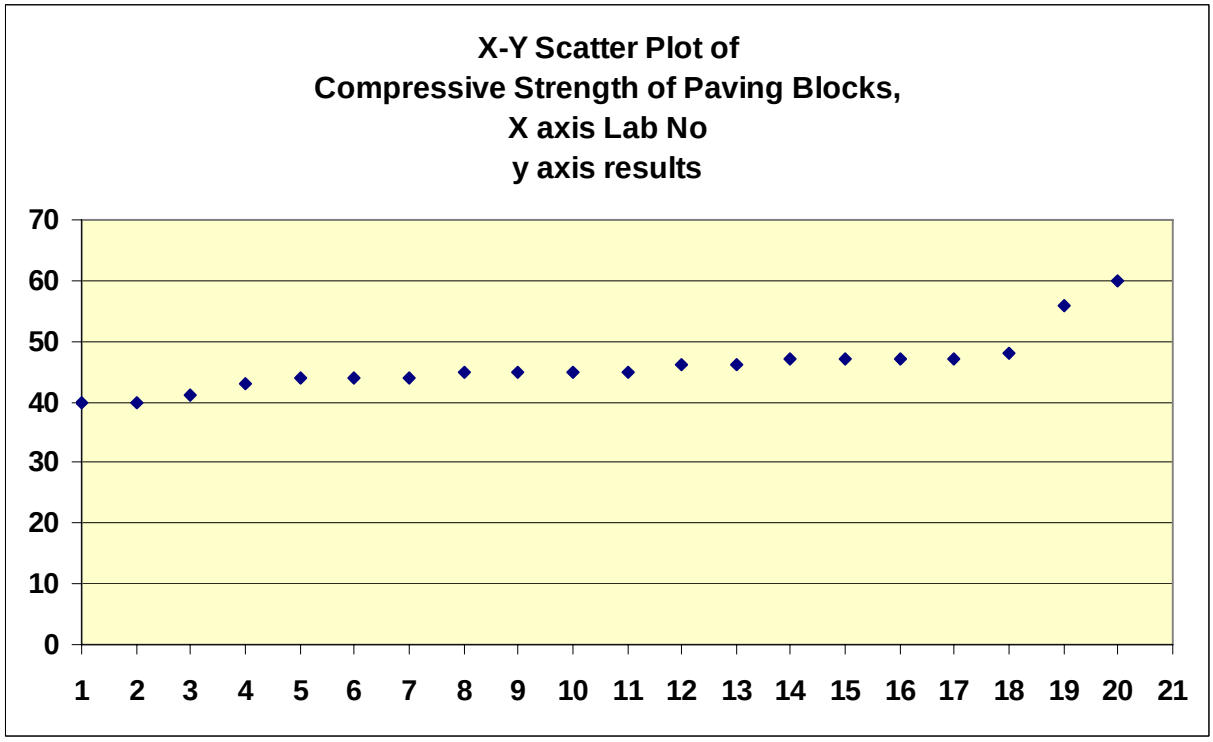
6.3 Identifying outliers

Table (1) shows that lab14 and lab18 " z-scores are exceeding +/-3 (i.e. 3.753 for lab14 and 5.143 for lab18) and thus, both results are considered outliers as per clause [5.3].

6.4 Charts

The charts described in clause [5.4] are shown below:


Z-Score of Compressive Strength of Paving Blocks,
6
5
4
3
2
1
0
-1
-2
-3
-4
-5
-6
Lab 16
Lab 8
Lab 3
Lab 6
Lab 10
Lab 12
Lab 19
Lab 13
Lab 21
Lab 4
Lab 9
Lab 11
Lab 7
Lab 15
Lab 17
Lab 2
Lab 5
Lab 1
Lab 14
Lab 18


X-Y Scatter Plot of
Compressive Strength of Paving Blocks,
X axis Lab No
y axis results
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21

7. References

7.1 Guide to NATA Proficiency Testing, 2004 Version 1.

7.2 ISO/ IEC 17011: General requirements for bodies providing assessment and accreditation of CABs.

7.3 ISO/ IEC 17025: General requirements for the competence of testing and calibration laboratories.

7.4 ILAC-G13: Guidance for the requirements for the competence of providers of proficiency testing schemes.

7.5 ISO/ IEC Guide 43: Proficiency Testing by Inter laboratory Comparisons- Part 1: Development and Operation of Proficiency Testing Schemes. Proficiency Testing by Inter laboratory Comparisons- Part 2: Selection and use of proficiency testing schemes by laboratory accreditation bodies.

7.6 ISO 13528: Statistical Methods for Use in Proficiency Testing by Inter-laboratory Comparisons.

7.7 Pure Appl. Chem. Vol. 78, No.1, pp. 145-196, 2006: The International Harmonized Protocol for the Proficiency Testing of Analytical Chemistry Laboratories (IUPAC Technical Report).

7.8 Nelac Institute, Environmental laboratory Sector, Volume 3, Proficiency Testing (PT) Provider Requirements, Interim Standard, December 18, 2007.

7.9 ILAC-P9:2005 ILAC Policy for Participation in National and International Proficiency Testing Activities.

7.10 DAC-Req-01 Accreditation Requirements of Dubai Accreditation Department.