

TCVN

TIÊU CHUẨN QUỐC GIA

TCVN 10860:2015

ISO 2602:1980

Xuất bản lần 1

**GIẢI THÍCH THỐNG KÊ KẾT QUẢ THỬ -
ƯỚC LƯỢNG TRUNG BÌNH - KHOẢNG TIN CẬY**

Statistical interpretation of test results - Estimation of the mean - Confidence interval

HÀ NỘI - 2015

Mục lục

	Trang
Lời nói đầu.....	4
Lời giới thiệu.....	5
1 Phạm vi áp dụng.....	7
2 Tài liệu viện dẫn.....	8
3 Thuật ngữ, định nghĩa và ký hiệu.....	8
4 Ước lượng trung bình.....	8
4.1 Trường hợp kết quả không được nhóm thành các lớp.....	8
4.2 Trường hợp kết quả được nhóm thành các lớp.....	8
5 Khoảng tin cậy cho trung bình.....	9
5.1 Ước lượng độ lệch chuẩn.....	9
5.2 Khoảng tin cậy cho trung bình.....	10
6 Trình bày kết quả.....	11
Phụ lục A (quy định) Khoảng tin cậy cho trung bình từ độ rộng mẫu.....	13

Lời nói đầu

TCVN 10860:2015 hoàn toàn tương đương với ISO 2602:1980;

TCVN 10860:2015 do Ban kỹ thuật tiêu chuẩn quốc gia TCVN/TC 69
Ứng dụng các phương pháp thống kê biên soạn, Tổng cục Tiêu chuẩn
Đo lường Chất lượng đề nghị, Bộ Khoa học và Công nghệ công bố.

Lời giới thiệu

Phạm vi của tiêu chuẩn này được giới hạn ở vấn đề riêng biệt. Nó chỉ liên quan đến việc ước lượng trung bình của tổng thể chuẩn trên cơ sở một loạt các thử nghiệm áp dụng với mẫu ngẫu nhiên cá thể lấy từ tổng thể này, đồng thời tiêu chuẩn chỉ đề cập đến trường hợp chưa biết phương sai của tổng thể. Tiêu chuẩn này không đề cập đến việc tính toán khoảng bao gồm ít nhất một tỷ lệ cho trước của tổng thể (giới hạn dung sai thống kê), với xác suất cố định.

Trong ISO 2854 đề cập đến các vấn đề sau đây (bao gồm cả vấn đề được đề cập trong tiêu chuẩn này):

- ước lượng trung bình và hiệu hai trung bình (đã biết hoặc chưa biết phương sai);
- so sánh trung bình với một giá trị đã cho và so sánh hai trung bình với nhau (phương sai đã biết hoặc chưa biết nhưng bằng nhau);
- ước lượng phương sai và tỷ số của hai phương sai;
- so sánh phương sai với giá trị đã cho và so sánh hai phương sai với nhau.

Phương pháp kiểm nghiệm thường được đưa ra cho nhiều phép xác định được tiến hành:

- trên cùng một cá thể (khi thử không phá hủy);
- với các tỷ lệ khác biệt của một sản phẩm có độ đồng nhất cao (ví dụ: chất lỏng);
- trên các cá thể riêng biệt được lấy mẫu từ một tổng thể với lượng biến động nhất định.

Trong hai trường hợp đầu, sai lệch giữa các kết quả thu được chỉ phụ thuộc vào độ lặp lại của phương pháp. Trong trường hợp thứ ba, sai lệch còn phụ thuộc vào độ biến động của bản thân sản phẩm.

Xử lý thống kê các kết quả cho phép tính toán khoảng chứa trung bình của tổng thể các kết quả thu được từ rất nhiều phép xác định, tiến hành trong cùng điều kiện, với một xác suất nhất định. Trong trường hợp cá thể có độ biến động, tiêu chuẩn này giả định rằng các cá thể được tiến hành xác định cấu thành một mẫu ngẫu nhiên từ tổng thể ban đầu và có thể được coi là độc lập.

Khoảng được tính toán như vậy gọi là khoảng tin cậy cho trung bình. Kèm theo nó là mức tin cậy (đôi khi gọi là hệ số tin cậy), là xác suất, thường được biểu thị bằng phần trăm, mà khoảng này chứa trung bình của tổng thể. Trong tiêu chuẩn này chỉ đưa ra các mức 95 % và 99 %.

Giải thích thống kê kết quả thử - Ước lượng trung bình - Khoảng tin cậy

Statistical interpretation of test results - Estimation of the mean - Confidence intervals

1 Phạm vi áp dụng

Tiêu chuẩn này quy định việc xử lý thống kê các kết quả thử nghiệm cần thiết để tính khoảng tin cậy cho trung bình của tổng thể.

Kết quả thử được biểu thị bằng các phép đo một đặc trưng liên tục. Tiêu chuẩn này không đề cập đến thử nghiệm đặc trưng định tính (ví dụ: có hay không có một tính chất, số khuyết tật, ...)

Phân bố xác suất được lấy làm mô hình toán học cho toàn bộ tổng thể là phân bố chuẩn, đối với phân bố này các thông số, trung bình m và độ lệch chuẩn σ là chưa biết.

Giả định về tính chuẩn được thỏa mãn rất rộng: phân bố của các kết quả thu được trong điều kiện thử nghiệm thường là phân bố chuẩn hoặc gần với phân bố chuẩn.

Tuy nhiên, có thể hữu ích khi kiểm tra hiệu lực của giả định về tính chuẩn bằng phương pháp thích hợp.

Việc tính toán có thể được đơn giản hóa bằng cách thay đổi kết quả gốc hoặc đơn vị của kết quả thử nhưng sẽ có rủi ro nếu làm tròn các kết quả này.

Không được loại bỏ bất kỳ quan trắc nào hoặc thực hiện hiệu chỉnh cho các quan trắc đáng ngờ mà không nêu rõ lý giải trên nền tảng thực nghiệm, kỹ thuật hoặc bằng chứng khác.

Phương pháp thử có thể có sai số hệ thống, việc xác định sai số này không được đề cập ở đây. Tuy nhiên, cần chú ý là sự tồn tại của sai số này có thể làm mất hiệu lực của phương pháp. Đặc biệt, nếu có độ chệch không nghi ngờ thì việc tăng cỡ mẫu n không ảnh hưởng đến độ chệch. Các phương pháp đề cập trong ISO 2854 có thể hữu ích trong những trường hợp nhất định để nhận biết sai số hệ thống.

2 Tài liệu viện dẫn

Các tài liệu viện dẫn sau rất cần thiết cho việc áp dụng tiêu chuẩn này. Đối với các tài liệu viện dẫn ghi năm công bố thì áp dụng phiên bản được nêu. Đối với các tài liệu viện dẫn không ghi năm công bố thì áp dụng phiên bản mới nhất, bao gồm cả các sửa đổi, bổ sung (nếu có).

TCVN 8244 (ISO 3534), Thống kê học – Từ vựng và ký hiệu

ISO 2854, *Statistical treatment of data – Problems of estimation and tests of means and variances* (Xử lý dữ liệu thống kê – Vấn đề ước lượng và kiểm nghiệm trung bình và phương sai)

3 Thuật ngữ, định nghĩa và ký hiệu

Tiêu chuẩn này sử dụng các thuật ngữ, định nghĩa và ký hiệu nêu trong TCVN 8244 (ISO 3534).

4 Ước lượng trung bình

4.1 Trường hợp kết quả không được nhóm thành các lớp

Sau khi loại bỏ mọi kết quả nghi ngờ, loạt gồm n phép đo x_i (trong đó $i = 1, 2, 3, \dots, n$), một số trong số này có thể cùng giá trị.

Trung bình m của phân bố chuẩn đang xét được ước lượng bằng trung bình cộng của n kết quả:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

4.2 Trường hợp kết quả được nhóm thành các lớp

Khi số lượng kết quả đủ lớn (ví dụ, trên 50) thì sẽ có lợi khi nhóm chúng thành các lớp có cùng độ rộng. Trong các trường hợp nhất định, kết quả cũng có thể thu được trực tiếp đã nhóm thành các lớp.

Tần số của lớp thứ i , tức là số các kết quả trong lớp i , được ký hiệu là n_i .

Số lớp được ký hiệu là k , ta có:

$$n_k = \sum_{i=1}^k n_i$$

Điểm giữa của lớp i được ký hiệu là y_i . Khi đó, trung bình m được ước lượng bằng trung bình có trọng số của tất cả các điểm giữa của các lớp:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^k n_i y_i$$

5 Khoảng tin cậy cho trung bình

Khoảng tin cậy cho trung bình tổng thể được tính từ các ước lượng của trung bình và ước lượng độ lệch chuẩn.

Phương pháp khác để tính khoảng tin cậy bằng việc sử dụng độ rộng mẫu được nêu trong Phụ lục A.

5.1 Ước lượng độ lệch chuẩn

5.1.1 Trường hợp kết quả không được nhóm thành các lớp

Ước lượng của độ lệch chuẩn σ , tính từ bình phương độ lệch so với trung bình cộng, được cho bằng công thức:

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$$

trong đó

x_i là giá trị của phép đo thứ i ($i = 1, 2, 3, \dots, n$);

n là tổng số các phép đo;

$\bar{x}$ là trung bình cộng của n phép đo, tính như nêu trong 4.1.

Để dễ tính toán, khuyến nghị sử dụng công thức sau:

$$s = \sqrt{\frac{1}{n-1} \left[\sum_{i=1}^n x_i^2 - \frac{1}{n} \left(\sum_{i=1}^n x_i \right)^2 \right]}$$

5.1.2 Trường hợp kết quả được nhóm thành các lớp

Trong trường hợp kết quả được nhóm thành các lớp, công thức cho ước lượng độ lệch chuẩn là

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^k n_i (y_i - \bar{y})^2}$$

Để dễ tính toán, khuyến nghị sử dụng công thức sau:

$$s = \sqrt{\frac{1}{n-1} \left[\sum_{i=1}^k n_i y_i^2 - \frac{1}{n} \left(\sum_{i=1}^k n_i y_i \right)^2 \right]}$$

trong đó

y_i là điểm giữa của lớp thứ i ($i = 1, 2, 3, \dots, k$);

k là tổng số các lớp;

n là tổng số các phép đo;

$\bar{y}$ là trung bình có trọng số của tất cả các điểm giữa của các lớp, tính như nêu trong 4.2.

Trong trường hợp kết quả được nhóm, giá trị tính được s cần được hiệu chỉnh ("hiệu chỉnh Sheppard"). Vì việc hiệu chỉnh này có tầm quan trọng thứ hai nên nó không được đề cập ở đây.

5.2 Khoảng tin cậy cho trung bình

Đối với mức tin cậy được chọn (95 % hoặc 99 %), theo trường hợp cụ thể, phải xác định khoảng tin cậy hai phía hoặc một phía.

5.2.1 Khoảng tin cậy hai phía

Khoảng tin cậy hai phía cho trung bình tổng thể được xác định bằng bất đẳng thức kép sau đây:

a) ở mức tin cậy 95 %:

$$\bar{x} - \frac{t_{0,975}}{\sqrt{n}} s < m < \bar{x} + \frac{t_{0,975}}{\sqrt{n}} s$$

b) ở mức tin cậy 99 %:

$$\bar{x} - \frac{t_{0,995}}{\sqrt{n}} s < m < \bar{x} + \frac{t_{0,995}}{\sqrt{n}} s$$

5.2.2 Khoảng tin cậy một phía

Khoảng tin cậy một phía cho trung bình tổng thể được xác định bằng một trong hai bất đẳng thức sau đây:

a) ở mức tin cậy 95 %:

$$m < \bar{x} + \frac{t_{0,95}}{\sqrt{n}} s$$

hoặc

$$m > \bar{x} - \frac{t_{0,95}}{\sqrt{n}} s$$

b) ở mức tin cậy 99 %:

$$m < \bar{x} + \frac{t_{0,99}}{\sqrt{n}} s$$

hoặc

$$m > \bar{x} - \frac{t_{0,99}}{\sqrt{n}} s$$

với $\bar{x}$, nếu cần, được thay bằng $\bar{y}$, trong trường hợp kết quả được nhóm thành lớp.

Giá trị $t_{0,975}$, $t_{0,995}$, $t_{0,95}$, $t_{0,99}$ là các giá trị có phân bố t Student với $v = n + 1$ bậc tự do.

Các giá trị này được cho trong Bảng 1.

Bảng này cũng đưa ra các giá trị của tỷ số

$$\frac{t_{0,975}}{\sqrt{n}}, \frac{t_{0,995}}{\sqrt{n}}, \frac{t_{0,95}}{\sqrt{n}}, \frac{t_{0,99}}{\sqrt{n}}$$

Khi giá trị của n lớn hơn 60, ưu tiên để tính giá trị của t bằng nội suy tuyến tính từ $\frac{120}{n}$ sử dụng Bảng 2.

VÍ DỤ:

$$n = 250$$

$$\frac{120}{n} = 0,48$$

$$t_{0,995} = 2,576 + 0,48 + (2,617 - 2,576) = 2,596$$

6 Trình bày kết quả

6.1 Đưa ra biểu thức trung bình theo 4.1 hoặc 4.2.

6.2 Biểu thị khoảng tin cậy dưới dạng bất đẳng thức kép của 5.2.1 hoặc một trong hai bất đẳng thức của 5.2.2, nêu rõ mức tin cậy (95 % hay 99 %). Chỉ ra số kết quả bị loại bỏ do nghi ngờ và lý do loại bỏ.

Bảng 1 – Giá trị của $t_{1-\alpha}$ và tỷ số $t_{1-\alpha}/\sqrt{n}$

n	Mức tin cậy Trường hợp hai phía		Mức tin cậy Trường hợp một phía		n	Mức tin cậy Trường hợp hai phía		Mức tin cậy Trường hợp một phía	
	95 %	99 %	95 %	99 %		95 %	99 %	95 %	99 %
n	$t_{0,975}$	$t_{0,995}$	$t_{0,95}$	$t_{0,99}$	n	$\frac{t_{0,975}}{\sqrt{n}}$	$\frac{t_{0,995}}{\sqrt{n}}$	$\frac{t_{0,95}}{\sqrt{n}}$	$\frac{t_{0,99}}{\sqrt{n}}$
2	12,71	63,66	6,314	31,82	2	8,985	15,013	4,465	22,501
3	4,303	9,925	2,920	6,965	3	2,484	5,730	1,686	4,021
4	3,182	5,841	2,353	4,541	4	1,591	2,920	1,177	2,270
5	2,776	4,604	2,132	3,747	5	1,242	2,059	0,953	1,676
6	2,571	4,032	2,015	3,365	6	1,049	1,646	0,823	1,374
7	2,447	3,707	1,943	3,143	7	0,925	1,401	0,734	1,188
8	2,365	3,499	1,895	2,998	8	0,836	1,237	0,670	1,060
9	2,306	3,355	1,860	2,896	9	0,769	1,118	0,620	0,966
10	2,262	3,250	1,833	2,821	10	0,715	1,026	0,580	0,892
11	2,228	3,169	1,812	2,764	11	0,672	0,956	0,546	0,833
12	2,201	3,106	1,796	2,718	12	0,635	0,897	0,518	0,785
13	2,179	3,055	1,782	2,681	13	0,604	0,847	0,494	0,744
14	2,160	3,012	1,771	2,650	14	0,577	0,805	0,473	0,708
15	2,145	2,977	1,761	2,624	15	0,554	0,769	0,455	0,668
16	2,131	2,947	1,753	2,602	16	0,533	0,737	0,438	0,651
17	2,120	2,921	1,746	2,583	17	0,514	0,708	0,423	0,627
18	2,110	2,898	1,740	2,567	18	0,497	0,683	0,410	0,605
19	2,101	2,878	1,734	2,552	19	0,482	0,660	0,398	0,586
20	2,093	2,861	1,729	2,539	20	0,468	0,640	0,387	0,568
21	2,086	2,845	1,725	2,528	21	0,455	0,621	0,376	0,552
22	2,080	2,831	1,721	2,518	22	0,443	0,604	0,367	0,537
23	2,074	2,819	1,717	2,508	23	0,432	0,588	0,358	0,523
24	2,069	2,807	1,714	2,500	24	0,422	0,573	0,350	0,510
25	2,064	2,797	1,711	2,492	25	0,413	0,559	0,342	0,498
26	2,060	2,787	1,708	2,485	26	0,404	0,547	0,335	0,487
27	2,056	2,779	1,706	2,479	27	0,396	0,535	0,328	0,477
28	2,052	2,771	1,703	2,473	28	0,388	0,524	0,322	0,467
29	2,048	2,763	1,701	2,467	29	0,380	0,513	0,316	0,458
30	2,045	2,756	1,699	2,462	30	0,373	0,503	0,310	0,449
40	2,024	2,707	1,682	2,430	40	0,320	0,428	0,266	0,384
50	2,008	2,680	1,675	2,404	50	0,284	0,379	0,237	0,340
60	2,000	2,664	1,673	2,393	60	0,258	0,344	0,216	0,309

Bảng 2 – Bảng nội suy tuyến tính khi n lớn hơn 60

n	$\frac{120}{n}$	$t_{0,975}$	$t_{0,995}$	$t_{0,95}$	$t_{0,99}$
60	2	2,000	2,664	1,673	2,393
120	1	1,988	2,617	1,658	2,358
∞	0	1,960	2,576	1,645	2,326

Phụ lục A

(quy định)

Khoảng tin cậy cho trung bình theo độ rộng mẫu

Nếu các phép đo được xếp theo thứ tự không giảm về độ lớn, sao cho $x_1 \leq x_2 \leq \dots \leq x_n$ khi đó $w = x_n - x_1$ được gọi là độ rộng mẫu. Vấn giả định rằng tổng thể có phân bố chuẩn, có thể xác định khoảng tin cậy cho trung bình tổng thể từ độ rộng mẫu khi số phép đo ít, tức là 12 hoặc ít hơn. Thực tế tính toán này có ưu điểm là nhanh hơn; nhược điểm là thường dẫn đến khoảng tin cậy rộng hơn và nhạy hơn với sai lệch so với dạng chuẩn giả định của các quan trắc.

Khoảng tin cậy hai phía

Khoảng tin cậy hai phía cho trung bình tổng thể được xác định bằng bất đẳng thức kép sau đây:

a) ở mức tin cậy 95 %:

$$\bar{x} - q_{0,975} w < m < \bar{x} + q_{0,975} w$$

b) ở mức tin cậy 99 %:

$$\bar{x} - q_{0,995} w < m < \bar{x} + q_{0,995} w$$

Khoảng tin cậy một phía

Khoảng tin cậy một phía cho trung bình tổng thể được xác định bằng một trong hai bất đẳng thức sau đây:

a) ở mức tin cậy 95 %:

$$m < \bar{x} + q_{0,95} w$$

hoặc

$$m > \bar{x} - q_{0,95} w$$

b) ở mức tin cậy 99 %:

$$m < \bar{x} + q_{0,99} w$$

hoặc

$$m > \bar{x} - q_{0,99} w$$

Các hệ số $q_{0,975}$, $q_{0,995}$, $q_{0,95}$, $q_{0,99}$ được cho trong Bảng A.1.

Bảng A.1 – Các hệ số $q_{0,975}$, $q_{0,995}$, $q_{0,95}$, $q_{0,99}$

	Mức tin cậy Trường hợp hai phía		Mức tin cậy Trường hợp một phía	
	95 %	99 %	95 %	99 %
n	$q_{0,975}$	$q_{0,995}$	$q_{0,95}$	$q_{0,99}$
2	6,353	31,828	3,157	15,910
3	1,304	3,008	0,885	2,111
4	0,717	1,316	0,529	1,023
5	0,507	0,843	0,388	0,685
6	0,399	0,628	0,312	0,523
7	0,333	0,507	0,263	0,429
8	0,288	0,429	0,230	0,366
9	0,255	0,374	0,205	0,322
10	0,230	0,333	0,186	0,288
11	0,210	0,302	0,170	0,262
12	0,194	0,277	0,158	0,241

Nguồn: E. Lord. Sử dụng độ rộng thay cho độ lệch chuẩn trong kiểm nghiệm t (*Biometrika*, Tập 34, 1947, pp. 41-67), với mục nhập $q_{0,95}$ ($n = 2$) được hiệu chỉnh.